Tumor-stroma derived gene expression patterns as prognosticators in breast cancer

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen

Fakultät der Universität Basel

Von

Michal Rajski

von Lodz, Polen

Basel, 2010

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Nancy Hynes
Prof. Dr. Gerhard Christofori
PD Dr. Martin Buess
Prof. Dr. Christoph Rochlitz
Prof. Dr. Ulrich Certa

(Mitglieder des Dissertationskomitees)

Basel, den 02.03.10

(Datum der Fakultätssitzung)

Prof. Dr. Eberhard Parlow

Dekanin/Dekan

Summary

The stroma, which embeds epithelial cells, plays a major role in the shaping and physiological regulation of animal organs such as the breast. Aberrant functioning of the stroma was found to support tumorigenesis and cancer progression. Among many crucial aspects of breast cancer biology two issues focused our attention; (I) the impact of mutual interactions between the cancer cells and the stroma, specifically the endothelial cells and (II) the impact of insulin like growth factor one (IGF-I) on the gene expression profiles of stromal fibroblasts, as an example of the influence of a growth factor on the stroma.

To address tumor-endothelial interaction we took advantage of an *in vitro* system mimicking the interaction of tumor and endothelial cells and explored gene expression changes using DNA microarrays. Our results suggest the interaction of endothelial cells and tumor cells that express the CD44+/CD24- signature indicative of stem cell-like cells, which are thought to have a low proliferative potential, might explain the unexpected and paradoxical association of the CD44+/CD24- signature with highly proliferative tumors that have an unfavorable prognosis. Furthermore, the gene expression signature induced in this system is of prognostic value in early stage and metastatic breast cancer.

Since malignant epithelial cells and tumor-associated stromal cells are under the influence of hormones and growth factors, we examined the effects of IGF-I on cancer cells and primary fibroblasts in parallel to assess concordant and discordant gene expression changes. Our results show that primary breast fibroblasts, breast carcinoma associated fibroblasts and primary lung fibroblasts respond to IGF-I stimulation with increased expression of genes related with proliferation, which is phenotypically followed by an increased growth rate. Comparing *in vitro* gene expression data with available *in vivo* data, we have shown that the evoked gene expression signatures are able to stratify patients into groups with significantly different outcome. We propose that this prognostic gene expression signature might also serve as a predictor for the effectiveness of an anti-IGF-I therapy, a new therapeutic strategy that is currently in phase III development.

Concluding, the data presented in this thesis underline the importance of the stroma, specifically the interaction between tumor and endothelial cells and the response of the stroma to stimulation with growth factors such as IGF-I. The gene expression changes in response to these interactions and stimulations carry prognostic information and might potentially be useful in clinical decision-making.

Introduction

Cancer epidemiology

Cancer is one of the major causes of death worldwide. The disease accounted for 7.9 million deaths (around 13% of all deaths) in 2007. This number is projected to continue rising, with an estimated 12 million deaths in 2030. Lung, stomach, colorectal, liver, and breast cancer cause most of cancer deaths each year [1]. Breast carcinoma, the most frequent type of cancer among women, was estimated to be responsible for 40,170 deaths in the United States in 2009 [2]. The high social impact of the disease implicates the need for more efficient treatments. In order to develop proficient treatments for the different forms of the disease, it is crucial to understand the biology of breast cancer, including role of stromal cells and tumor-stroma interactions in tumorigenesis and breast cancer progression.

Cell types and microenvironmental factors affecting breast cancer development and progression

Breast cancer is biologically and clinically a heterogeneous and multistage disease. It is mainly a malignancy of the mammary ductal or lobular epithelial cells (carcinomas), but can, rarely, affect the stromal tissue of the breast (sarcomas). At the initial stage hyperproliferative epithelial cells are contained entirely *in situ*, within the duct (ductal carcinoma *in situ* - DCIS) or lobule (lobular carcinoma *in situ*). Early breast DCIS lesions either do not evolve [3-4], or progress into invasive ductal carcinomas (IDC), when hyperproliferative cells break through the basement membrane surrounding the duct and invade into the stromal tissue of the breast. The process is schematically shown in figure 1.

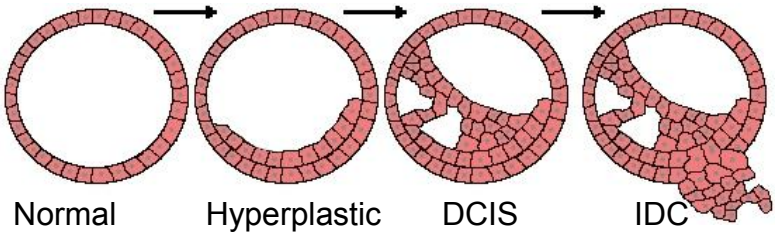


Figure 1. Schematic illustration of changes accompanying progression from healthy mammary duct to hyperplastic changes to ductal carcinoma *in situ* (DCIS) to invasive ductal carcinomas (IDC) (from the left to the right).

This stage is followed by multiple events including tumor vascularization and transition of epithelial cells into a mesenchymal state [5]. Vascularization facilitates tumor growth [6], and the epithelial to mesenchymal transition (EMT) allows cancer cells to spread. EMT is reversible and at the place of metastasis the cells shift back to the epithelial state and form tumors with the full characteristics of a primary tumor. As the changes (transition) are reversible they cannot be explained by genetic alterations, which suggests that there must be other regulatory components present, among them the tumor microenvironment [7]. In breast cancer, the population of malignant cells is heterogeneous as it consists of different cell clones and cell types (cancer stem cells and differentiated cancer cells) [8]. Moreover, the tumor bulk is comprised of cells of multiple origins, including fibroblasts, endothelial, myoepithelial, various immune cells and the extracellular matrix (ECM) [9]. All of the components intermingle and interact with many signaling loops implicated either through soluble factors [10-11] or involving direct interactions [12]. Additionally, the tumor is located inside the living organism and it is under constant, systemic influence of hormones, growth factors that actively modify the tumor microenvironment [13-16]. The general idea is shown in figure 2.

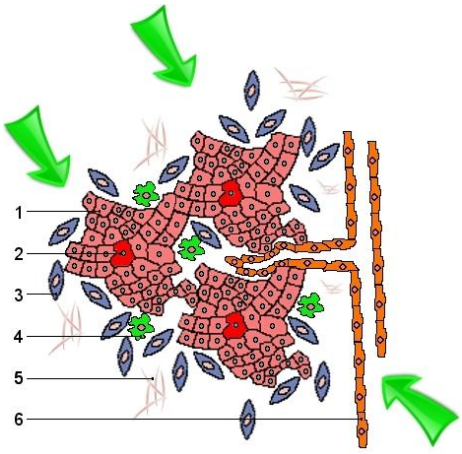


Figure 2. Schematic image illustrating the complexity of the breast cancer tumor bulk. Two populations of cancer cells (1) differentiated ancestors and (2) cancer stem cells are marked. Different cells building the tumor stroma like; (3) fibroblasts, (4) immune system cells e.g. macrophages (6) endothelial cells and (5) components of extra cellular matrix are visualized. Green arrows symbolize the systemic influence of hormones and growth factors acting on the cancer cells and on the carcinoma associated stroma.

Different populations of cancer cells within tumor

With the success of molecular *in vitro* studies, cancer research was focusing predominantly on the epithelial cells. Tumors have been understood as a homogenous population of highly proliferative malignant cells. Currently it is known that the malignant epithelial cell fraction contains a subset of cells with stem cell properties and their differentiated progenies.

There are many reasons of malignant transformation and many cells are regarded to be cancer cell's ancestors. Particular attention is given to somatic stem cells, as they are long living and therefore hypothesized to be able to accumulate multiple mutations since they are exposed to damaging agents or an unfavorable environment for a long time. The transition of a somatic stem cell into a cancer stem cell (CSC) is plausible because of the propensity of self-renewal and multiple divisions that might be the cause of fixation and propagation of alterations in their genome [17]. It was recently shown that tumor protein 53 (p53), a well known tumor suppressor gene, regulates the polarity of cell division in mammary stem cells and that the loss of p53 favors symmetric divisions of stem cells, contributing to an increased amount of stem cells [18]. An increased number of symmetrically dividing stem cells increases the amount of target cells for transformation. After the transformation, cells with mutated p53 continue to divide symmetrically, increasing the amount of cancer stem cells within the tumor. The idea is schematically illustrated in figure 3. The concept of cancer stem cells, first proven in acute myeloid leukemia [19], is nowadays widely confirmed for numerous types of cancer including breast cancer [8, 20-22]. It is hypothesized that the cancer stem cell fraction is likely to be responsible for the resistance to chemotherapy [23] e.g. colon cancer stem cells were proven to produce and utilize IL-4 to protect themselves from apoptosis [24]. There are many markers that are suspected to define cancer stem cell including different combinations of CD24, CD29, epithelial specific antigen (ESA), CD44, CD49f, CD133, stem cell antigen one (Sca1), (reviewed in [25]) but the unambiguous marker constellations, characterizing a cancer stem cell, is unknown. Since the

markers are not yet precisely described detailed characterization of CSC, leading to effective targeting is still missing.

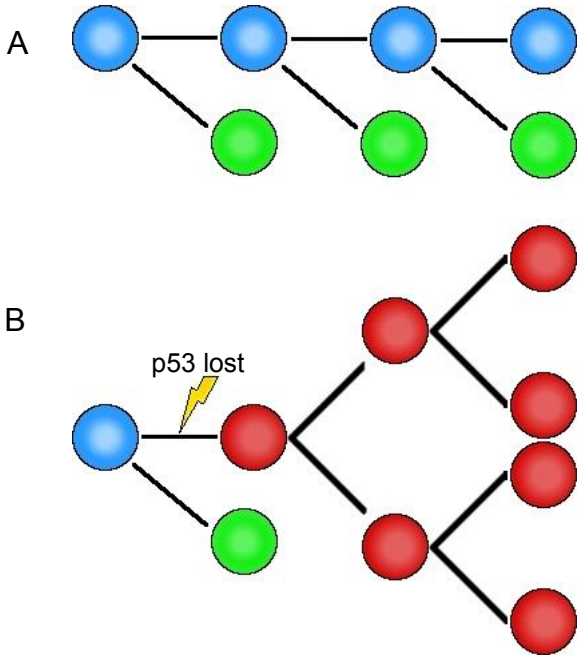


Figure 3. Schema illustrating the impact of p53 mutation on the amount of mammary stem cells. Part A depicts physiological situation with symmetric division of the mammary stem cell (blue) into 2 daughter cells: the renewed stem cell (blue) and its differentiated progeny (green). Part B depicts symmetric divisions of mammary stem cells with a lost of p53 (red). Increased amount of mammary stem cells increases the amount of target cells for transformation and later facilitates the amount of cancer stem cells within the tumor.

Microenvironment affects tumorigenesis and cancer progression

It has been shown in animal xenograft models that injection of purified malignant epithelial cells results in formation of histologically-complex tumors, with as many as 80% of the cells being stromal cells [26]. From developmental biology it is known that diverse types of cells building the mammary gland collaborate closely in its development, and the types of cells involved are highly conserved. All of them cooperate to establish the whole mammary gland and the mammary stem cell niche [27]. Similarly, during the neoplastic process, all cell types intermingle and interact. The cells building the tumor stroma are distinct from their normal counterparts [28], they have modified characteristics leading to active formation of the tumor microenvironment [29-32]. The fact that the microenvironment affects the efficiency of tumor formation, growth, invasiveness and metastatic potential was shown in many examples. A typical example of the

microenvironment leading to cancer is chronic inflammation, caused by *Helicobacter pylori* infection in stomach and leading to stomach cancer and hepatitis C infection of the liver, leading to chronic inflammation, proceeding to liver cirrhosis and, not infrequently, liver cancer [33]. In other experiments, injection of non-transformed mammary epithelial cells into irradiated mammary stromal fat pads resulted in increased tumor growth when compared to those injected into contralateral, non irradiated mammary fat pads [34]. The authors concluded that irradiated stromal cells alter the microenvironment and may lead to tumor promotion as illustrated in experiments with irradiated stroma. Moreover, malignant cells can exist within normal tissues but be restrained by normal milieu cues, what was concluded from experiments showing that similar patterns of mutations are found in both, normal epithelial tissue adjacent to tumor tissue and tumor tissue itself [35-36]. The authors concluded that, the mutation that initiates the carcinoma occurs in the epithelium, but events that promote tumor progression involve the stroma. In some cases, the trigger for neoplastic progression is speculated to come from signals within the stromal microenvironment [29]. In humans, early breast lesions are often present, however only a limited quantity of them progress towards malignancy. In the late eighties, Nielsen *et al.* analyzed 110 breast tissue autopsy samples and found that 32% of the patients displayed hyperplastic lesions, 27% had atypical ductal hyperplasia and 18% were DCIS positive [4]. Similar data were presented by Alpers *et al.* [3]. The high numbers of lesions not progressing into invasive carcinomas suggests that other mechanisms leading to malignancy must be involved including a “tumor prone” microenvironment as shown by Hu *et al.* [37]. Animal tumor models have shown that fibroblasts over expressing hepatocyte growth factor (HGF) or transforming growth factor *beta* (TGF-β) are able to induce formation of tumors at diverse sites including stomach and prostate [38-39]. Additionally, cancer-associated stromal cells are different from their corresponding normal analogues in terms of gene expression patterns already at a preinvasive (DCIS) tumor stage [28]. Summarizing, the described data suggest that microenvironment and stromal cells comprising for it are crucial not only in

tumor invasion and metastasis, but also in the earlier steps of breast tumorigenesis.

Fibroblasts in tumor progression

One of the main pieces of evidence for the role of fibroblasts in tumor progression comes from mouse xenograft models. Co-injection of tumor cells with stromal cells (fibroblasts) induces a faster manifestation of bigger tumors [40-41]. It has been hypothesized that the delay in single-cell type xenografts is due to the time required for the cancer cells to recruit supportive cells to the tumor (e.g. from the bone marrow [42]), to form the complex structure and create signaling networks [43]. The instigating role of the stroma was extensively investigated and many factors taking part in the cross-talk between tumor and stroma cells were characterized. Fibroblast secreted protein-1 (FSP1; mts1) is a protein secreted by both fibroblasts and cancer cells making the environment more favorable for tumor progression as it regulates angiogenesis and inflammation and is responsible for metastatic cancer progression [44]. Moreover, tumors forming after co-injection of carcinoma cells with *Fsp1-/-* fibroblasts into *Fsp1 -/-* animals had significantly decreased numbers of infiltrating macrophages, smooth muscle actin-expressing myofibroblasts, and CD31-positive endothelial cells, compared to tumors developing after coinjection of tumor cells with *Fsp1+/+* fibroblasts on the same mouse background [44]. Lack of CAFs expressing FSP1 in this system resulted in lower percentage of tumor formation and those formed did not metastasize at all. Other studies prove that CXCL12 (SDF-1, stromal cell-derived factor 1), expressed by fibroblasts, stimulates cancer cell proliferation by acting through C-X-C chemokine receptor type 4 (CXCR4) [11]. CXCL12 was proposed to stimulate metastasis to lung and lymph nodes through high expression of CXCL12 at these organs, resulting with homing of CXCR4 positive cancer cells to these organs. Additionally, CXCL12 was shown to have an impact on angiogenesis as it is involved in recruitment of endothelial cell progenitors to the growing tumor [11]. Since CXCL12 is strongly chemotactic for lymphocytes [45] and all of the aforementioned data come from immuno-compromised mice, it was not possible to assess the effects of CXCL12

mediated interactions between CAFs, leukocytes and tumor cells. Another molecule illustrating the role of the stroma in tumorgenesis and progression is type I collagen. Its increased amount is responsible for a high mammographic breast density, which correlates with an increased risk of sporadic breast cancer [46]. Carcinoma associated fibroblasts (CAF) are the cells responsible for the change of extracellular matrix composition to one with increased amounts of collagens (desmoplastic response) [26]. Therefore, fibroblasts producing increased collagen levels might be involved in sporadic breast cancer incidence [29]. Stromal fibroblasts also have an impact on tumor stroma composition by expression of different metalloproteinases, namely metalloproteinase 13 (MMP-13), which is expressed by CAF-like cells in human breast cancer [47]. *In vivo*, breast cancer cells can stimulate fibroblasts to secrete MMP-13 [48]. MMP-13 acts on the proteins building the ECM and modulates signaling pathways from the ECM and modulates the bioavailability of growth factors. Metalloproteinases are important in EMT transition and increased invasiveness of the breast cancer [49-50] as they help to break the basement membrane and release cancer cells, which is one of the most reliable signs of poor prognosis in most carcinoma systems.

Concluding, fibroblasts, in particular the CAFs, actively support tumor cells and modify the tumor environment to make it more advantageous for tumor progression.

Endothelial cells in tumor progression

Endothelial cells, together with pericytes, form the tumor neovasculature, which supplies tumor cells with nutrients and oxygen, and removes waste and carbon dioxide. Primary tumors without vasculature are typically small, dormant nodules of tissue whose volume remains constant by a balance of cell proliferation and cell death [51]. Further growth of the tumor mass induces hypoxic conditions in the center of tumor bulk that induces expression of vascular endothelial growth factor-A (VEGF-A) and subsequently tumor vascularization [52]. Enhanced angiogenesis is associated with an increased risk of metastasis and poor

prognosis [53]. Although highly important, vascularization is not the only role of endothelial cells, as they also serve as an important source of cancer growth regulation (e.g. for liver cancer cells *in vitro* [54]). Endothelial cells are also involved in the establishment of the cancer stem cell niche and metastatic spread of tumor cells into distant organs [55]. Tumor cell interaction with the endothelium during hematogenous dissemination, and the following interaction with endothelium and subendothelial matrix constitute the most crucial factors in determining the organ preference of metastasis. Cell surface adhesion molecules (i.e., integrins, cadherins, immunoglobulins and selectins) and many other unspecified molecules, mediate tumor-host endothelium interactions [31]. Selection of a place of metastasis is not the only process involving the attachment of tumor cells to the endothelium. The attachment as well provides the necessary anchorage that prevents anoikis and facilitates cancer cell proliferation [30]. Targeting tumor angiogenesis was shown to be an effective therapeutic option [52]. Nonetheless, both in preclinical and clinical settings, the benefits are of short duration and are followed by a restoration of tumor growth and progression, even with increased aggressiveness [56]. Pre-existing or acquired resistance to anti-angiogenic therapy might be mediated by factors that act through local paracrine loops between tumor cells and endothelial stromal cells in manner similar to CAFs that secrete platelet-derived growth factor C (PDGF-C), which in turn stimulates tumor angiogenesis [57] in answer to hypoxia induced by anti-VEGF therapy.

Other stromal cells in tumor progression

In addition to fibroblasts and endothelial cells, tumor stroma involves numerous cell types including immune cells, bone marrow derived cells and several factors constituting ECM.

The host defense system comprises numerous types of cells and factors that, in the context of tumor biology, should concomitantly work to eradicate a tumor. Often the system is deformed, resulting in a benefit for the tumor i.e. as shown for inflammatory cells promoting tumor angiogenesis. Monocytes (at the tumor

site macrophages), lymphocytes and neutrophils are recruited to tumor stroma, where they release the variety of factors that alter cellular behavior. Classical examples are VEGF, HGF, metalloproteinase 2 (MMP-2) and interleukin 8 (IL-8) released by macrophages and neutorophils that affect endothelial cells contributing to tumor progression [32].

Using *in vivo* experiments Karnoub *et al.* [10] demonstrated that mixing bone-marrow-derived human mesenchymal stem cells (hMSCs) with weakly metastatic breast cancer cells and injecting them subcutaneously in the mouse resulted in tumors with an increased metastatic potential when compared to breast cancer cells injected alone. The presence of cancer cells stimulated bone-marrow-derived hMSCs to secrete chemokine (C-C motif) ligand 5 (CCL5, RANTES), which acted back on cancer cells in a paracrine manner [10]. In other studies adult human mesenchymal stem cells enhanced breast tumorigenesis and promoted hormone independence. Co-injection of MCF-7 cells and hMSCs in mice induced hormone independence and increased proliferation and additionally sensitized tumors to estradiol. These tumors had probably as well increased estrogen signaling [58].

As a consequence of the studies focusing almost exclusively on cancer cells, nearly all of the currently used cancer therapeutic agents target the cancer cells themselves, which, due to their inherent genomic instability, frequently acquire therapeutic resistance [28, 59]. Due to the fact that stromal cells are thought to be genetically more stable and less heterogeneous than tumor cells, they seem to be an interesting drug target. If aiming to search for a target within the stroma we have to keep in mind that it mutually co-evolves together with the tumor [60]. Ma *et al.* recently confirmed this by conducting a comparative analysis of global gene expression changes in the stromal and epithelial compartments during breast cancer progression from normal to pre-invasive to invasive ductal carcinoma. They concluded that tumor-associated stroma undergoes extensive gene expression changes during cancer progression, to a level similar to the malignant epithelium [61]. From this point of view a solution might be targeting a specific interaction, signaling loop, between cancer cells and stroma cells that

remains stable during cancer progression. A successful example of such an approach is VEGF blocking [52].

Despite these convincing data implicating a role of the tumor microenvironment and stromal cells in breast tumorigenesis, our understanding of the genes mediating cellular interactions and paracrine regulatory circuits among various cell types in normal and cancerous breast tissue and their role in breast tumorigenesis is limited [28]. In analogy to our work exploring tumor fibroblast interaction [12] we planned to take advantage of an *in vitro* system mimicking the tumor – endothelial interaction, to explore the effects of heterotypic interactions on global gene expression and describe the pathways involved in signaling between tumor cells and endothelium.

Systemic influence of hormones and growth factors on cancer initiation and progression

Another important factor regulating tumor progression is the involvement of hormones. The serum levels of some of them are positively correlated to breast cancer incidence [14, 62-63], e.g. sex hormones like estrogens. Tumor progression is also dependent on growth factors e.g. transforming growth factor *beta* (TGF-β) signaling has multiple, context-dependent roles in human cancers ranging from arrest of cell growth to induction of migration, stimulation of epithelial to mesenchymal transition and tumor progression [13]. In addition to direct effects on tumor cells, hormones and growth factors might modulate tumor-stroma interactions. Most carcinoma-associated fibroblasts, known to support cancer growth, express *alpha* smooth muscle actin (αSMA) and have an increase in contractility, which indicates that the majority of these cells may be myofibroblasts. TGF-β is capable of inducing the transformation of fibroblasts into myofibroblasts. Furthermore, genetic removal of TGF-β receptor type two in stromal fibroblasts resulted in carcinoma of adjacent epithelial cell populations [13].

Estrogens are female sex hormones with a critical impact on reproduction and sexual functioning. A clear impact of estrogens on cancer development and

progression was shown in numerous models including an observation from 1896, when G.T. Beatson described that bilateral oophorectomy resulted in the remission of breast cancer in premenopausal women [64]. Amongst many hazard factors, an excess of estrogens was proven to increase the risk of breast cancer. Case-control studies (New York University Women's Health Study and the Ormoni e Dieta nell'Eziologia dei Tumori study) indicated that increased levels of estrone, estradiol and bioavailable estradiol, as well as their androgenic precursors, may be associated with a 4- to 6-fold increase in the risk of postmenopausal breast cancer [63]. Moreover, estrogens administrated during hormone replacement therapy (estrogen plus progestin pills) increase the risk of invasive breast cancer. The Women's Health Initiative Estrogen-plus-Progestin Study showed that, after 5 years of follow-up, women receiving the hormones had a 24 percent increase in breast cancer risk compared with women in the placebo group [65]. From 51 to 82% of human breast cancers are estrogen receptor alpha (ER) positive and ER status is considered an important prognostic factor. The proportion of ER positive tumors increases with patient age [66]. Estrogens promote the development of mammary cancer in rodents and exert both direct and indirect proliferative effects on human breast-cancer *in vitro*. Direct tumor-initiating effects involve enzymes and proteins involved in nucleic acid synthesis and through the activation of oncogenes. Indirect effects may occur through the stimulation of prolactin secretion and the production of growth factors (e.g., transforming growth factor *alpha* and epidermal growth factor) and non–growth-factor peptides (e.g. plasminogen activators) [14]. Estrogen signaling in the tumor bulk is not only based on endocrine sources. Carcinoma-associated fibroblasts actively induce local estrogen synthesis, which contributes to carcinogenesis and progression [67]. The dependence of ER-positive cancer cells on estrogen provides a treatment option and drugs interfering with estrogen receptors are currently used in clinics.

Prolactin (PRL) is a polypeptide hormone released from the anterior pituitary gland that stimulates milk production after childbirth. PRL can promote epithelial cells proliferation and survival, increase cell motility and support tumor cell

vascularization [16]. *In vitro,* it increases the growth of normal and malignant breast cells [68-70]. In animal models, it is important for mammary epithelial development and its administration has been shown to increase the rate of mammary tumor formation [71]. Animal data suggest that PRL can induce estrogen ER negative (ER-) tumors or rarely estrogen receptor positive (ER+) tumors in mouse models [16]. Ingram *et al.* found that in human, prolactin levels higher than the median value in control subjects were associated with a more than two-fold increase in the risk of breast cancer [15], which was confirmed in a prospective study in postmenopausal women by Hankinson [62]. In addition, a modest positive association between prolactin and breast cancer risk was found among predominately premenopausal women [72].

Similar to the two aforementioned hormones, other hormones and growth factors influence development and progression of breast cancer. Insulin is a growth-promoting hormone that is involved in the pathogenesis of various malignancies that may act as a growth hormone through regulation cell proliferation, differentiation and apoptosis. Insulin mediates its effects by binding to the transmembrane insulin receptor (IR) what leads to activation different mitogen-activated protein (MAP) kinases (MAPK) and PI3K (Phosphoinositide 3-kinases) pathway [73-74]. Insulin enhances the production and mitogenic activity of growth factors such as epidermal growth factor and insulin-like growth factors what is suspected to lead to pathological cell proliferation [75]. Hyperinsulinemia (excess levels of circulating insulin in the blood) associated with insulin resistance, a typical feature of type II *Diabetes mellitus*, is one of the risk factors in the development of various malignancies, including breast, pancreas, liver, colon, bladder, and oral cavity cancer [74]. Furthermore, population studies have shown that women with a new diagnosis of type two diabetes have a significantly larger likelihood of a prior diagnosis with breast cancer. This association allows speculation that the elevated insulin levels during pre-diabetes phase are promoting breast cancer development and progression [76]. Most important, breast tumors over-express IR and increased IR expression correlates with poor survival [77]. Moreover, insulin is able to increase ER expression and insulin and

estradiol synergistically accelerate breast cancer cell proliferation [77]. In addition, insulin has indirect effects on breast cancer progression including the ability to simulate aromatase activity thereby increasing the levels of bioavailable estradiol [78]. The similarity of IR signaling pathway with insulin-like growth factor one receptor (IGF-IR) signaling pathway allows speculation that the tumorigenic and tumor supporting actions of both are similar or somehow intermingled. The harmful impact of IGF-I in breast cancer development and progression is well documented and IGF-I-blocking strategies are currently under extensive investigation [79]. On the contrary to tumor cells, the impact of the IGF-I on the cells *in vitro* representing the stromal fraction of the tumor bulk, is an essentially unexplored area. Fibroblasts are the most abundant cells in the tumor stroma mass mediating large parts of the paracrine signaling between epithelial cells and stroma [26]. Stromal cells respond to IGF–I stimulation with increased proliferation, as shown for fibroblasts [80-81] and microvascular endothelial cells [82]. In the project presented in the second part of this thesis I sought to characterize gene expression changes induced by IGF-I stimulation on primary breast fibroblasts and tried to check if the gene expression changes accompanying IGF-I stimulation are of prognostic value.

DNA Microarrays as a tool to profile gene expression changes

DNA microarray technology has evolved from Southern blotting and enables the simultaneous measurement of expression levels of thousands of genes. The first microarray prototypes were used for gene expression profiling in 1987 [83] and the use of miniaturized microarrays for gene expression profiling was first reported in 1995 [84]. Since the first application, the technology evolved rapidly and currently various approaches (single or double color) and formats based on different concepts are in use. The most common format is a glass slide with many DNA fragments, known as probes, covalently attached. Since a single microarray slide contains thousands of probes, a microarray experiment can accomplish thousands of genetic tests in parallel. Presently, microarrays might be applied for various applications ranging from gene expression profiling, comparative genomic hybridization (CGH), alternative splicing detection to single

nucleotide polymorphism (SNP) detection. The performance of the microarrays is astonishing. An example is the Roche NimbleGen microarray that allows for simultaneous gene expression analysis of 12 patient samples on a single, high resolution slide (2um pixel resolution). Each patient sample is checked with 135000 probes for complete coverage of the human transcriptome. In addition to the compact format, the high resolution provides better signal-to-noise performance. Furthermore, simultaneous processing of samples grants high level of inter-array reproducibility (r^2=0.99). Although the technology seems expensive, as it requires a high resolution scanner and special hybridization chambers, the final cost per sample, especially for industrial purposes, is reasonable [85].

HEEBO-arrays

In my work, I applied the two-color Human Exonic Evidence Based Oligonucleotide microarrays (HEEBO). The general experimental design of two color microarray format is visualized in figure 4. HEEBO microarrays consist of 44,544 70mer probes, which include (a) constitutive exonic probes (30,718), (b) alternatively spliced / skipped exonic probes (8,441), (c) non-coding RNA probes (196), (d) BCR / TCR Genic / Regional Probes (372), (e) other Probes (843) and (f) controls. HEEBO microarrays were produced at the Stanford Functional Genomic Facility (Stanford, USA). Complete details regarding the clones on the arrays may be found at Stanford functional genomics facility website [86].

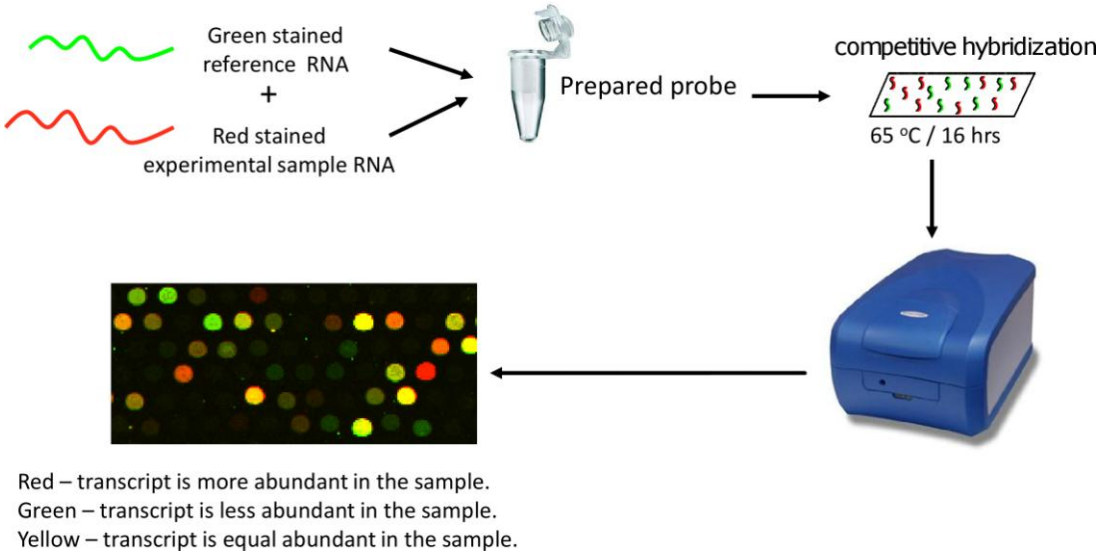


Figure 4. Schematic illustration of the experimental procedure of the two-color microarray approach.

Gene expression patterns as a tool to systematize cancer

The progress of molecular profiling of solid tumors began in 2000 when Perou *et al.* proposed that phenotypic diversity of breast tumors might be mirrored by the diversity in gene expression patterns and that the systematic investigation of gene expression patterns might improve molecular taxonomy of breast cancers. The study, based on 42 patients and cell lines, had shown that breast tumors greatly differ in their gene expression patterns and the differences allow formulating a taxonomy of breast cancers. The division of the tumors based on expression levels of the so called “intrinsic” gene set allowed to formulate four groups: ER+/luminal-like; basal-like; *Erb-B2* and normal breast [87]. The main caveat of this study is the limited amount of samples that did not allow distinguishing cancer subtypes with full resolution. Nonetheless, in addition to the first molecular breast cancers taxonomy, there were two striking conclusions gained from this work. First, so called “estrogen receptor negative” carcinomas covered at least two distinct tumor subtypes (basal-like and *Erb-B2*). Second, the molecular program of gene expression patterns of primary tumors was recapitulated in its metastases. In the follow-up study, Sorlie *et al.* increased the amount of specimens and provided a classification of 78 locally-advanced breast carcinomas [88]. The obtained classification confirmed the one previously proposed. The novel finding was that formerly defined ER+/luminal-like group could be divided into three subgroups: luminal A, luminal C and luminal B. In addition, the tumor gene expression characteristics were correlated with clinical outcome. Correlation of the five most distinctive groups with clinical data showed that all of them are unique in terms of overall-survival and relapse-free survival. The follow-up study, based on 115 samples, confirmed classification to five molecular subtypes of breast cancer and validated it within two independent datasets [89]. Additional inclusion of patients with mutations of BRCA1 (breast cancer 1, early onset) and BRCA2 (breast cancer type 2 susceptibility protein) genes revealed the similarity of this group of tumors to the basal subtype breast cancers that are ER negative and carry poor prognosis.

Concluding, molecular profiling of breast carcinomas allowed a more detailed

classification than classic pathological measurements and demonstrated that transcriptional programs in the tumor cells and the underlying genetic alternations are major determinants of the tumorgenic potential and ultimately the clinical outcome of the disease [89]. Molecular profiling of tumors was done also for other malignances including lung adenocarcinomas [90-91], lymphomas, liver cancer, ovarian cancers and soft tissue tumors [89, 92-94]. Since most of these analyses were performed using bulk tissue samples that are composed of multiple cell types or purified tumor epithelial cells, the specific contribution of epithelial and stromal cells to these tumor classifiers and prognostic signatures remained unknown [28].

Gene expression signatures as a tool to predict patients' survival

Parallel to the molecular profiling of tumors, an extensive work was conducted to assess if microarray derived gene expression patterns, called signatures, might be used as a tool to predict patients' survival. The correlation of signatures with clinical characteristics was conducted by many researchers [87-88, 95-102] (reviewed in [103]). The best known, "70-genes" signature derived by Laura van t' Veer *et al.* [104] in a supervised analysis of early stage breast cancers, is a potent gene expression pattern outperforming all known clinical predictors. It was created to distinguish genes strongly predictive of a short interval to distant metastasis for lymph-node positive and negative patients. The validity of the signature was confirmed in an independent study by Van de Vijer *et al.* within the dataset containing 295 samples [105]. Based on the signature, a small custom-made microarray – "MammaPrint" was created. Results of the validation demonstrated that microarray technology could be used as a reliable diagnostic tool [106]. Further validation in an independent group of 307 patients (137 events after a median follow-up of 13.6 years, no systemic therapy) from five European cancer centers has shown that the "70-genes" signature is able to add independent information for clinicopathologic risk assessment for patients with early breast cancer [107]. In another prospectively conducted study, including 427 patients in various cancer centers, the use of MammaPrint in combination

with standard clinical guidelines led to altered adjuvant treatment recommendations in 26% of patients [103]. Currently, the MINDACT (Microarray In Node-negative and 1 to 3 positive lymph node Disease may Avoid ChemoTherapy) trial is being conducted. The trial is a multicentre, prospective, phase III randomized study comparing the 70-gene expression signature with a common clinical-pathological prognostic tool (Adjuvant! Online) selecting patients for adjuvant chemotherapy in node-negative breast cancer. Up to 14th of October 2009, 4114 patients entered the screening phase and 2264 were enrolled. The final goal is to enroll 6000 patients. The primary objective of the MINDACT trial is to confirm that patients with a “low risk” molecular prognosis and “high risk” clinical prognosis can be safely spared chemotherapy without affecting distant metastasis free survival (DMFS). MINDACT has several secondary objectives including the identification and validation of novel gene expression signatures predicting clinical response to therapies used (chemotherapy and endocrine therapy). Furthermore, an estimation of the efficacy of chemotherapy in terms of disease free survival, DMFS and overall survival in the two subgroups where the clinical-pathological prognosis and the molecular prognosis are discordant will also be performed. The project also aims to set up several tissue bank resources (RNA, tumor tissue, serum) for future translational research studies in both genomics and proteomics [108-109]. The results are still unknown but potentially, the signature might provide a method to tailor adjuvant systemic treatment and reduce the costs and side effects of unnecessary treatment. The limitation of the method is that MammaPrint test requires fresh tissue that has to be shipped to company producing MammaPrint for analysis.

This limitation is overcome in an alternative prognostic test; OncoType DX created by Genomic Health [110] that might be processed from the formalin-fixed, paraffin-embedded tissue sample. Also this test quantifies the likelihood of disease recurrence in women with early-stage breast cancer. It is based on quantitative reverse transcription polymerase chain reaction (qRT-PCR) and assesses the expression of a panel of 21-genes to determine the probability of disease recurrence. It consists of 16 cancer genes and five reference genes used

to normalize the expression of the cancer genes. Oncotype DX was evaluated in clinical trials involving over 3,300 patients with the conclusion that it not only quantifies the likelihood of breast cancer recurrence in women with node-negative, estrogen receptor-positive breast cancer, but also predicts the magnitude of chemotherapy benefit [111-113].

The possibility of connecting gene expression patterns with clinical data resulted in a large number of publications showing correlations between biological *in vitro* derived signatures and clinical characteristics of tumors sharing similar gene expression patterns. Some of them were able to stratify patients into groups with different survival [12, 87-89, 98, 100-107, 111-123]. Comparing all of those predictive signatures raised two questions; first, why is there nearly no gene overlap between the "70-genes" signature and other predictive signatures and the second; are the predictions derived from these gene signatures concordant for individual patient samples. The first topic was addressed by Ein-Dor *et al.* [124]. The authors focused at van't Veer's dataset [104] and repeated the methodology of the "70-genes" signature creation, changing only the initial step – the subset of patients used to create the predictor. The results showed that the set of predictive genes is not unique and it strongly depends on the subset of patients used to create the predictor. In other words, there are many prognostic signatures with similar predictive power even within one dataset. Furthermore results by Ein-Dor *et al.* showed that there are many genes correlated with survival and the differences between these correlations are small. This allows hypothesizing that every single gene caries a bit of information and microarrays, as a tool picturing a complex image of cells tested, have to be analyzed in a "context" manner in which the general deregulation of a pathway or a process must be an informative event and not a single gene expression fluctuation. The second question, if predictions derived from these gene sets are concordant for individual samples, was addressed by Fan *et al.* [125]. Within a single data set of 295 samples [105] the authors applied five gene-expression-based models: intrinsic subtypes [89], 70-gene profile [104], wound signature [117], recurrence score [112], and the two-gene ratio [126] and compared the results for single

patients and found that most models had high rates of concordance in their outcome predictions for the individual patients.

Another important factor that has to be taken into account is that the majority of breast cancer patients today receive some form of postoperative treatment (radiation, endocrine and/or chemotherapy) that influences the clinical course and significance of prognostic factors. It is possible that creating the prognostic classifier (signature) from patients who are treated with different therapies with significant impact on survival might introduce a bias in prognostic capacities of the signature. Although the resulting predictive classifier may accurately describe the sample set used for its development, it may fail when applied to an independent test set containing patients which still differ in terms of outcome [127]. This issue was kept in mind when validation of the "70-genes" signature by Buyse *et al.* [107], was conducted. Similarly, another signature presented in a study by Wang *et al.* was created using gene expression of frozen tumor samples from 286 lymph-node-negative patients who had not received adjuvant systemic treatment. The authors identified a 76-gene signature that was highly informative in identifying patients who developed distant metastases within 5 years in pre- and postmenopausal women [122].

An exciting possibility that might be achieved with microarray-based technology is creating a single microarray test covering prognosis, ER and HER2 status and sensibility to various treatment approaches. Invention and application of such custom microarray would provide clinicians with all required information for individual future treatment options shortly after tumor resection.

Stroma and tumor-stroma interaction derived gene expression signatures as a tool to predict patients' survival

Gene expression profiling of tumors is a field that developed with similar trends to those in cancer biology. Gene expression profiling studies started with a focus on epithelial cells that widen to tumor stroma, finally to notice the importance of single tumor stromal cell types that are now extensively studied.

One of the first gene expression signatures that might be named "stromal" was a signature obtained by West *et al.* [123] from soft tissue sarcomas that recapitulate features of normal connective tissue; solitary fibrous tumors – STF (derived from fibroblasts) and desmoid-type fibromatosis – DTF (derived from deep soft tissue fibroblasts). These two sarcoma types have very distinct gene expression profiles including differential expression of extracellular matrix and growth factors genes. By immunochemistry, cells positive for gene markers specific for DTF were found in reactive tissues like scar and inflammatory granulation tissue and cells positive for STF markers were found in normal tissue. A signature of 786 genes that was able to distinguish the STF from DTF was defined. These patterns of expression were also present in breast cancer and were able to stratify breast cancer patients into two groups with significantly different outcomes. The data obtained by West *et al.* were evaluated in a follow-up study, involving four more independent datasets [114]. In these four datasets, the authors identified a core set of 66 DTF-associated genes that were coordinately expressed in a subset of 25-35% of breast cancer patients. Patients with tumors representing high levels of this signature tend to have a better prognosis. In addition, the authors evaluated the protein expression of a single DTF core protein - secreted protein, acidic, cysteine-rich (SPARC), and showed it to be frequently expressed in the tumor stroma and absent in non-neoplastic tissue. Furthermore, breast cancer patients with strong stromal expression of SPARC showed a trend to increased survival. Concluding, the signature representing genetic alterations in tumors originating from fibroblasts was a valid prognosticator for breast cancer patients. Yet, one has to remember that the signature was not derived from stroma cells itself but malignant mesenchymal cells and thus may represent specific sarcoma cells gene expression profile and not solely a cancer stroma gene imprint.

The first signature derived from stromal cells, namely fibroblasts, was created by Chang *et al.* [117], based on the observation of histological similarities between a healing wound and the tumor microenvironment resembling "normal wound healing gone awry" [128]. In the study, fibroblasts from ten anatomical sites were

stimulated with serum and evoked gene expression patterns were analyzed. The “wound signature”, containing 512 stereotype genes up-regulated in fibroblasts upon serum stimulation was postulated to reflect the role of fibroblasts in wound healing and was hypothesized to provide a molecular estimate for the presence of the wound-healing process in human cancers. The signature was able to distinguish between patients with significantly different outcome for different types of human cancers such as locally advance breast carcinoma, lung adenocarcinomas (two datasets) and gastric carcinomas. However, it failed to distinguish patients with diffuse large B-cell lymphoma, medulloblastoma, and glioblastoma multiforme. Immunochemical staining revealed that LOXL2 (lysyl oxidase-like 2), an extracellular matrix remodeling enzyme present in the signature, was exclusively expressed by peritumoral fibroblasts around invasive carcinomas (45 of 106 tumors examined). LOXL2 primarily produced by fibroblasts may act on endothelial cells during tissue remodeling. Collectively, the data presented in this study show that the gene expression profile of fibroblasts stimulated by serum in human cancers represents a multicellular program in which all of the cells building the “tumor-bulk” actively participate, and the gene expression signature evoked in fibroblasts might be a valid predictor in breast cancer. In a follow-up study [116], the authors proved that the “wound signature” is able to stratify breast cancer patients in an independent early breast cancer dataset into two groups with significantly different prognoses. Moreover the signature was able to split breast cancer patients into groups that do or do not require adjuvant chemotherapy more efficient than National Institutes of Health [129] or St. Gallen [130] criteria, suggesting that treatment decision based on “wound signature” might be beneficial for patients [116], sparing some women the toxic chemotherapy.

It is worth noting that the “wound signature” [117], the molecular classification of breast cancers [89] and “70-genes” signature [104] proved that gene expression patterns are a tool to classify the tumors and predict their course either by direct survival fitting (“70-genes” signature) or by *in vitro* modeling (“wound signature”). Both of the signatures classified tumors into coherent and internally consistent

groups, and where the signatures diverged, the combined information gave improved risk stratification compared to individual signatures [116].

An important contribution to gene expression profiling of tumor stroma was done by Allinen *et al.* [28] who used cell-type specific markers to isolate epithelial and stromal cell populations from normal breast tissue, DCIS and invasive breast carcinomas and analyzed the gene expression of single cell types with serial analysis of gene expression (SAGE). The gene expression data showed that not only epithelial cells, but also stromal cells are different from corresponding counterparts already at the pre-invasive DCIS stage. The study revealed high expression levels of genes coding CXCL12 and chemokine (C-X-C motif) ligand 14 (CXCL14) at both DCIS and invasive stages by myoepithelial cells and myofibroblasts. Both chemokines are known to stimulate cancer cell growth through CXCR4 [131]. In addition, array comparative genomic hybridization (aCGH) analysis revealed that genetic changes like chromosomal gains or losses are limited only to tumor epithelial cells. This study provided a comprehensive analysis of gene expression of normal breast, DCIS and invasive breast carcinomas and showed that stromal cells associated with tumors are different from corresponding normal cells. Since CXCL12 and 14 produced by myoepithelial cells are able to stimulate cancer growth, they represent interesting drug targets.

Gene expression profiling of the breast cancer derived fibroblasts in invasive breast cancer was conducted by Singer *et al.* [132]. They compared gene expression patterns of cultured stromal fibroblasts obtained from malignant tissue of 10 women with gene expression patterns of cultured normal breast tissue fibroblasts of 10 women with benign breast disorders. Out of 2400 genes examined with DNA microarrays, 135 were more than two fold up-regulated and 105 more than 0.5 fold down-regulated in fibroblasts from malignant tissue. The majority of genes with increased expression levels were encoding tumor-promoting cytokines (e.g. colony stimulating factor 1), transcription factors (HYL) and cell-matrix associated proteins (cadhepsin L). Osteopontin, recently shown to be involved in mobilizing stromal cell precursors from the bone marrow into the

circulation, thereby making them available for tumors [42], was significantly up-regulated in breast cancer derived fibroblasts. An important observation taken from this study is that fibroblasts retain their distinctive gene expression pattern in the absence of tumor epithelial cells in cell culture. These authors hypothesized that the carcinoma-associated fibroblasts represent a specific fibroblast subpopulation which favors tumor growth [132].

Recent findings by Casey *et al.* [133] confirmed that cancer-associated stroma (fibroblasts) differs from the normal mammary stroma in gene expression of genes mostly involved in ECM remodeling (MMP-1, MMP-13, FAP, TLL2), as well as SPARC, POSTN (periostin) and TGF-β. Moreover, using microarrays, the author profiled all laser capture microdissected (LCM) stromal components of tumor and compared it to corresponding parts of normal breast, concluding that breast cancer invasion proceeds through the acquisition of a motile phenotype in tumor epithelial cells and a reactive phenotype in cancer associated fibroblasts.

Gene expression profiling of the endothelial cells in invasive breast cancer was conducted by Parker *et al.* [53]. Purified endothelial cells from either normal mammoplasty or primary breast tumors were used to generate a SAGE database of gene expression changes accompanying vascular proliferation in invasive breast cancer. Invasive breast cancer vasculature, in contrast to normal endothelium, expressed ECM and surface proteins characteristic of proliferating and migrating vasculature cells. Furthermore, compared to other types of cancer (colon cancer), invasive breast cancer endothelial cells induced high expression levels of two transcription factors SNAIL1 (Snail 1 Drosophila homolog) and HEYL (hairy/enhancer-of-split). HEYL was shown to induce proliferation and attenuate apoptosis of primary endothelial cells *in vitro*. In addition, PRL3 (Protein-tyrosine phosphatase type IVa, member 3), another intracellular protein, was shown to be predominantly expressed by endothelial cells of invasive breast cancer and being able to enhance the migration of endothelial cells *in vitro* [53]. The data presented by Parker *et al.* confirmed that tumor-associated endothelial cells differ from corresponding normal counterparts in terms of gene expression

and some of the dissimilarities are specific only for breast cancer.

These studies, as well as a tumor microenvironment characterization provided by Allinen [28], give an insight in carcinoma-associated stroma biology and proves that they are co-conspirators, rather than innocent bystanders, of cancer progression. Although the authors did not check the prognostic capacities of the signatures in human cancer gene expression databases, I speculate that some of them would be valid predictors as they cover genes involved in processes crucial in instigating tumor growth. Entirely tumor stroma-derived microarray gene expression pattern used as a prognostic marker for clinical outcome was published by Finak *et al.* [120]. Laser microdissected stromal compartments from 53 primary breast tumors were gene expression profiled together with matching normal tissue. Genes differentially expressed between tumor stroma and normal stromal tissue were used to create a stroma derived prognostic predictor (SDPP). SDPP has shown increased accuracy with respect to previously published predictors [89, 106, 116, 119, 123], especially for HER-2 positive tumors. The efficiency of the signature in independent, whole-tumor derived gene expression datasets suggests that changes in breast tumor stroma have an essential role in disease progression and outcome and the information that it carries might be found in whole tumor bulk gene expression profiles. Moreover, the prognostic capacity independent of ER and HER-2 status allows speculation that the information carried by the SDPP identifies processes different from those associated with tumor clinical subtype, like specific receptor status. As mentioned above, the signature was created based on the entire tumor stroma. This has some advantages, such as including genes involved in immunological responses, angiogenesis and hypoxia. The disadvantage is as that using a mixture of the cells does not allow distinguishing the cells responsible for gene overexpression. Furthermore it does not reveal any signaling loops that may take place between the cells. However a plus of this predictive signature is its size (26 genes). It is smaller than the complementary SFT/DTF signature what makes it more applicable for diagnostic practice. The small size also allows speculation that the most important actions of the stroma are pictured without blurring the image with

redundant information. In addition, it reduces the chance of distribution of the patients to groups with different outcome based on genes that carry only little information. The strong predictive power suggests that all of the genes included in the signature are important. SDPP was derived from the whole stroma of 53 primary breast tumors, 50 of them represented invasive ductal carcinoma (IDC). As it is known that tumor-associated stroma undergoes extensive gene expression changes during cancer progression, to an extent similar to the malignant epithelium [61], it is possible that SDPP represents alterations of tumor stroma only at the IDC stage. This allows speculation that the signature composition could differ if other stages of tumor progression were included e.g. the stroma surrounding DICS when the basement membrane is intact.

A subtype of gene signatures derived from stromal components are gene expression profiles derived from co-culture studies. An example is the interferon response signature [12]. In experiments involving *ex-vivo* co-culture models, simulating tumor stroma interaction, the authors found that in the proximity to cancer cells the fibroblasts secreted type I interferons, which induced expression of the interferon response genes (IRG) in the tumor cells. Paralleling this model, immunohistochemical analysis of human breast cancer tissues showed that STAT1, the key transcriptional activator of the IRGs, and itself an IRG, was expressed in a subset of the cancers, with a striking pattern of elevated expression in the cancer cells in close proximity to the stroma [12]. The signature, induced by co-culture, was able to segregate 295 early-stage breast cancers into two groups with significantly different outcomes. Since fibroblasts from different body sites have different gene expression programs [134-135], a chance for differential interaction of primary CAFs with breast tumor cells exists. This limitation was partially overcome in a similar work conducted by Sato and colleagues for co-culture of pancreatic cancer cells and primary pancreatic fibroblasts. This analysis identified multiple genes as differentially expressed in co-culture as a consequence of mutual interaction, including genes associated with tumor invasion, metastasis and angiogenesis. *In vitro* the invasive potential of pancreatic cancer cells (CFPAC1) was increased when they were co-cultured

with primary pancreatic fibroblasts. Selective inhibition of cyclooxygenase 2, a gene found to be up-regulated in both cell types during co-culture, partially decreased the invasive properties of cancer cells [136]. Unfortunately, the authors did not correlate the obtained gene expression signature to clinical data so one cannot conclude if using site-matching fibroblasts influences the prognostic capacity.

When one is thinking about tumor expression profiling, it is important to remember that gene expression studies of tumors are based on samples containing a variety of cells [88-89, 104-105, 122]. In general, for cancer microarray profiling, biopsy samples containing less than 50% of cancer cells are excluded from the studies [137]. This reduces the amount of the samples that are profiled, thereby reducing the number of patients that can benefit from the gene expression profiling. It is crucial to consider the degree to which inclusion of stromal cells influences the outcome of tumor profiling studies. This issue was addressed by Roepman *et al.* [137]. The prognostic capacity of a previously formulated gene expression signature [138] varied depending on the cancer cell content in the biopsy. Interestingly not only low (<50%) but also too high (>90%) cancer cell content in gene expression profiled biopsy decreased the capacity to foresee metastasis in head and neck squamous cell carcinoma (HNSCCs) based on the signature. To analyze the problem, they created artificial samples with different percentages of LCM dissected stromal and tumor cells admixed and determined the gene expression profiles. With this approach it was possible to assess which of the genes were stroma and which were tumor specific. The results showed that 12% of genes in the previously established predictive signature are predominantly expressed in the stroma, 25% are specific for tumor cells and the rest of genes are equally expressed in the tumor and the stroma. In other words, the strength of the signature to foresee presence of metastasis in HNSCCs is associated with up-regulation of a set of genes specific for the stroma in the tumor bulk. This data underlines both the importance of the stromal component in gene expression profiles and in metastatic process, and shows that profiling of cancer cells only may introduce a bias and decrease the predictive

power of the gene expression signatures. Another important observation was made by Bryant *et al.* who quantified the proportion of lipopolysaccharide-induced differentially-expressed monocyte genes that could be measured in peripheral blood mononuclear cells, and determined the extent to which gene expression in the non-monocyte cell fraction diluted or obscured fold changes that could be detected in the cell mixture. They concluded that the effect of non-responding cells in a mixture obscures the detection of subtle changes in gene expression in an individual cell type. However, for studies in which only the most highly differentially expressed genes are of interest, separating and analyzing individual cells might be unnecessary as the mixture does not hide those genes [139]. Concluding, in the search for highly differentially-expressed genes, the screen done with microarray technology does not require purification to single cell type, as most deregulated genes aren't going to be hidden. On the other hand if one searches for subtle differences between normal and cancer-associated stroma, purification of the cells to a single cell type suspension seems to be crucial.

Tumor-stroma derived gene expression patterns as prognosticators in breast cancer

Concluding, tumor gene expression signatures are valid predictors in cancers including breast cancer. Signatures derived from carcinoma-associated cells, fibroblasts under the conditions mimicking wound healing, or tumor-stroma interactions are also valid predictors. Since the stroma plays a crucial role in cancer progression, new drugs targeting different aspects of tumor-stroma interactions or stroma itself are needed. To describe alterations taking place in either primary breast fibroblasts under the influence of growth factors or in tumor stroma interactions we sought to characterize global gene expression changes in both systems using microarray technology. Gene expression signatures derived from both *in vitro* systems will be validated in publically available gene expression datasets.

Results

Tumor-endothelial interaction links the CD44(+)/CD24(-) phenotype with poor prognosis in early-stage breast cancer

Buess M, Rajski M, Vogel-Durrer BM, Herrmann R, Rochlitz C.

Neoplasia. 2009 Oct;11(10):987-1002.

NEOPLASIA
www.neoplasia.com

Volume 11 Number 10 October 2009 pp. 987–1002 **987**

Tumor-Endothelial Interaction Links the CD44$^+$/CD24$^-$ Phenotype with Poor Prognosis in Early-Stage Breast Cancer[1]

Martin Buess*,†, Michal Rajski†, Brigitte M.L. Vogel-Durrer†, Richard Herrmann†,‡ and Christoph Rochlitz†,‡

*Medical Oncology, St. Claraspital, Basel, Switzerland; †Department of Biomedicine, University of Basel, Basel, Switzerland; ‡Department of Medicine, Division of Oncology, University of Basel, Basel, Switzerland

Abstract

MATERIALS AND METHODS: The genomic effects of tumor-endothelial interactions in cancer are not yet well characterized. To study this interaction in breast cancer, we set up an *ex vivo* coculture model with human benign and malignant breast epithelial cells with endothelial cells to determine the associated gene expression changes using DNA microarrays. *RESULTS:* The most prominent response to coculture was the induction of the M-phase cell cycle genes in a subset of breast cancer cocultures that were absent in cocultures with normal breast epithelial cells. In monoculture, tumor cells that contained the stem cell–like CD44$^+$/CD24$^-$ signature had a lower expression of the M-phase cell cycle genes than the CD44$^-$/CD24$^+$ cells, and in the CD44$^+$/CD24$^-$ cocultures, these genes were induced. Pretreatment gene expression profiles of early-stage breast cancers allowed evaluating *in vitro* effects *in vivo*. The expression of the gene set derived from the coculture provided a basis for the segregation of the tumors into two groups. In a univariate analysis, early-stage tumors with high expression levels (n = 137) of the M-phase cell cycle genes had a significantly lower metastasis-free survival rate (P = 1.8e − 5, 50% at 10 years) and overall survival rate (P = 5e − 9, 52% at 10 years) than tumors with low expression (n = 158; metastasis-free survival, 73%; overall survival, 84%). *CONCLUSIONS:* Our results suggest that the interaction of endothelial cells with tumor cells that express the CD44$^+$/CD24$^-$ signature, which indicates a low proliferative potential, might explain the unexpected and paradoxical association of the CD44$^+$/CD24$^-$ signature with highly proliferative tumors that have an unfavorable prognosis.

Neoplasia (2009) 11, 987–1002

Introduction

Tumor angiogenesis is a prerequisite for tumor progression and metastasis. It is a complex process that requires cooperative reciprocal interaction of tumor cells and endothelial cells [1–4] and, thereby, offers an attractive therapeutic target [5]. Clinical trials with antiangiogenic agents, such as bevacizumab, which is an antibody against vascular endothelial growth factor (VEGF), introduced these agents into clinical practice [6]. During the last several years, antiangiogenic therapies, in combination with conventional chemotherapeutic agents, have been established for different tumor types, such as colorectal cancer [7], non–small cell lung cancer [8], renal cell cancer [9], and breast cancer [10]. The average clinical benefit of these drugs, however, is relatively modest, and it is unclear which patients benefit the most. Improvements are likely to come from a more thorough understanding of the molecular and cellular mechanisms that govern tumor-endothelial cell interactions. Tumor angiogenesis involves a plethora of soluble and cellular components that interact in a process of mutual signaling [11]. This requires a coordinated expression of proangiogenic factors [12] and suppression of antiangiogenic factors [13], which leads to endothelial cell proliferation and migration and vessel formation. Although multiple single genes have been described in numerous reports to be involved in angiogenesis, such as

Address all correspondence to: Martin Buess, St. Claraspital, Kleinriehenstrasse 20, CH-4016 Basel, Switzerland. E-mail: Martin.Buess@claraspital.ch

[1]This work was supported in part by the Swiss National Science Foundation grant 320000-112794/1 and a grant from Oncosuisse.
Received 21 April 2009; Revised 13 July 2009; Accepted 13 July 2009

DOI 10.1593/neo.09670

growth factors [12,14], membrane-bound molecules [15], and extracellular matrix components [16], there are likely others that have remained unidentified. The interplay between the various factors and their combined effects in tumor angiogenesis, however, remains to be further characterized.

Carcinomas are not merely aggregates of malignant epithelial cells but are, in many respects, organlike structures that include host stromal cells, such as fibroblasts, adipocytes, inflammatory cells, and the cells that form the tumor vasculature, and the malignant cells themselves that intermingle and interact with all of these cell types [17]. During the last few years, there has been growing evidence that, besides the cellular processes within the tumor cells, a relevant contribution to tumor progression is provided by the cells of the tumor microenvironment [18]. On the molecular level, genome-scale gene expression studies of many different carcinomas have illustrated in detail the complexity of the tumors and the diversity of the associated non–epithelial cell types [19]. Inductive interactions between these different cell types can play not only a morphogenetic role but also an important mechanistic role in the pathogenesis and progression of malignancy. The endothelial cells have so far been mainly viewed in the context of vessel formation to improve the blood supply of the tumor. However, relatively little is known about the paracrine effects of these tumor-endothelial cell interactions. It was commonly thought that the formation of new vessels would mainly be important for the transport of nutrients and oxygen to the tumor cells and that interrupting this support is the key mechanism of antiangiogenic therapies. If we assume that, by the interruption of the vascular support, the tumor gets more hypoxic, it seems paradoxical that antiangiogenic therapies enhance the effects of chemotherapy and radiation. In the hypoxic environment, these therapies have usually been shown to be less effective [20]. However, the effects of these agents could be due, in good part, to the disruption of the paracrine tumor-promoting signaling that occurs as a result of the interaction of the cancer and endothelial cells. Such reciprocal inductive signaling has been well known from developmental biology and has again attracted special attention with the concept of cancer stem cells and their stem cell niche [21]. Therefore, characterizing heterotypic cell-cell interaction effects on a global gene expression scale might help to better understand the currently used antiangiogenic agents and eventually lead to the identification of novel targets that could be used to interrupt these paracrine stimulatory signaling pathways. This study specifically focuses on the interaction between breast cancer cells and endothelial cells to identify their reciprocal signaling effects.

Breast cancer is a heterogeneous disease, which implies that the tumor-endothelial cell interactions might also be diverse. Tumor-endothelial cell interactions are not yet well characterized on a genome-wide scale, and they have not been compared between different tumor subtypes. Toward this aim, we performed a systematic analysis of the interactions between well-characterized breast cancer cell lines and primary endothelial cells in coculture.

We have recently used the approach of *in vitro* coculture experiments to characterize heterotypic interaction effects with DNA microarrays to systematically describe the global-scale effects that the tumor-fibroblastic stroma interaction has on gene expression. We identified a strong induction of an interferon response by specific tumor cells in coculture with a diverse set of fibroblasts, which corresponded to a subset of breast cancers with an unfavorable prognosis *in vivo* [22].

In this study, we used breast cancer cell lines and endothelial cells for systematic coculture experiments, which allowed the interaction effects to be characterized on a global gene expression scale. Using this system, we investigated the following hypotheses:

1. The interaction of tumor and endothelial cells leads to changes in gene expression, which are important for angiogenesis and tumor progression. The gene expression programs that are involved provide hints for the signaling mechanisms that are involved.
2. The interaction of tumor and endothelial cells leads to the induction of gene expression signatures that are clinically relevant. These interaction effects might account for a significant proportion of the unexplained information in the gene expression data from tissue specimens. Given the evidence that interactions between cells can play critical roles in tumor progression, these data might be even more meaningful than prominent expression patterns, which are driven by the proportional representation of a given cell type in a tissue.

We have performed a systematic overview of heterotypic interaction effects of breast cancer and endothelial cells. The picture we obtained is complex, but our results suggest that the interaction of endothelial cells with a subset of $CD44^+/CD24^-$ breast cancer cell lines induces a signature of "tumor-endothelial cell–induced M-phase cell cycle" genes, which is associated with a worse outcome in human breast cancer.

Materials and Methods

Cell Culture

Human mammary epithelial cells (HMECs; Cambrex Bio Science Walkersville, Walkersville, MD) were expanded in mammary epithelial basal medium that was supplemented with bovine pituitary extract, human epidermal growth factor, insulin, and antibiotics (Clonetics, Cambrex Bio Science Walkersville). MCF-7, T47D, MDA-MB-231, SKBR-3, Hs578T, and BT549 (ATCC, Atlanta, GA) were propagated in Dulbecco's modified Eagle medium that was supplemented with 10% FBS (HyClone, Logan, UT), glutamine, 100 U/ml penicillin, and 100 μg/ml streptomycin (Gibco, Grand Island, NY). Human umbilical vein endothelial cells (HUVECs; ATCC) and human dermal microvascular endothelial cells (HDMECs; Cambrex Bio Science Walkersville) were expanded in endothelial basal medium 2 (EBM2; Cambrex Bio Science Rockland, Rockland, ME) that was supplemented with human epidermal growth factor, hydrocortisone, GA-1000 (gentamicin, amphotericin-B), 5% FBS, VEGF, human fibroblast growth factor-B (with heparin), R3-IGF, and ascorbic acid. For the coculture experiments, the cells were cultivated for 48 hours at an equal density of 50,000 cells/cm^2 (25,000 tumor cells/cm^2 and 25,000 endothelial cells/cm^2) in endothelial basal medium (Cambrex Bio Science Rockland) supplemented with 0.2% FBS. This medium served as a good universal medium for all the cells in the study.

Proliferation Assays

Direct cell counting. For cell counting, prestarved cells were plated in quadruplicate in 24-well plates at a density of 50,000 cells/cm^2. After 24 and 48 hours, the cells were trypsinized and resuspended in 0.2 ml of FACS buffer that contained 0.5% BSA and 2 mM EDTA in PBS. The total cell number was determined using a cell counter.

WST-1 assay. The proliferation reagent 4-[3-(4-iodophenyl)-2-(4-nitro-phenyl)-2*H*-5-tetrazolio]-1,3-benzene disulfonate (WST-1; Roche

Diagnostics GmbH, Roche Applied Science, Mannheim, Germany) was used according to the manufacturer's instructions. WST-1 is changed by mitochondrial enzymes of metabolically active cells to a colorful formazan that can be measured at a wavelength of 450 nm. The cell number was determined by comparison of the absorbance values to a standard cell dilution curve.

Comparison of HUVECs Proliferation in Response to Different Conditioned Medium

To obtain the conditioned medium, 10e6, Hs578T, MDA-MB-231, MCF-7, or HUVECs were extensively washed to avoid transfer of any stimuli from the regular cell growth medium. The cells were kept in 10 ml of EBM2 that contained 0.2% FBS for 24 hours. The medium was then aspirated and filtered through a 0.2-nm pore filter. In parallel, 3000 HUVECs per well were plated in a 96-well plate and starved for 24 hours in EBM2 that contained 0.2% FBS. For the stimulation experiments, HUVECs were washed once with PBS and incubated for 48 hours in 1:2 diluted conditioned medium in EBM2 that contained 0.2% FBS. EBM2 that contained 0.2% FBS was used as a negative control (vehicle medium), HUVEC culture supernatant that was diluted 1:2 in vehicle medium was used as an autologous medium control, and the full endothelial cell growth medium 2 that contained 5% FBS and all the supplements that were described above was used as a positive control. To determine the cell growth in response to stimulation with conditioned medium, the cells were stained with 7% WST-1 in low-serum Dulbecco's modified Eagle medium (0.2% FBS) for 1 hour at 37°C in 5% CO_2. The absorbance was measured with an ELISA reader at a wavelength of 450 nm. Values for each experimental condition were obtained by calculating the average of at least eight independent replicates.

Inhibition with the Blocking Antibody Bevacizumab

To block the stimulatory effects of tumor cell conditioned media, the medium was supplemented with 100 ng/ml bevacizumab (Avastin; Genentech/Roche). To control the effect of bevacizumab, an additional HUVEC culture was set up with a medium that was enriched with recombinant VEGF-A (5 ng/ml) and 5% FBS. The cell growth was determined as above, and the absolute absorbance values were compared.

To check if blocking is specific for bevacizumab, trastuzumab (Herceptin; Genentech/Roche) was included at equal concentrations (0, 360, 720 ng/ml) in the analysis. As a negative control, HUVECs that were treated with Hs578T cell culture supernatant were used. The cell growth was determined as described before, and the absolute absorbance values were compared.

RNA Isolation and Amplification

After discarding the culture medium and washing the cell layer once with PBS, total RNA was isolated by lysing the cells in the culture dish with RLT buffer (Qiagen, Valencia, CA) and extracting with the RNeasy Mini Kit (Qiagen). Five hundred nanograms of total RNA was amplified using the Message Amp II aRNA Kit (Ambion, Austin, TX). The amplification product was checked for integrity by electrophoresis in a 1% agarose gel in MOPS buffer.

Complementary DNA Microarrays and Hybridization

The human complementary DNA (cDNA) microarrays contained 40,700 elements, which represented 24,472 unique genes that were based on Unique Clusters. The arrays were produced at the Stanford Functional Genomic Facility. Complete details regarding the clones on the arrays may be found at: http://www.microarray.org/sfgf/jsp/servicesFrame.jsp#productionArrays.

cDNA produced from 6 μg of amplified RNA were hybridized to the array in a two-color comparative format, with the experimental samples labeled with one fluorophore (Cy5) and a reference pool of messenger RNA (Universal Human Reference mRNA; Stratagene, La Jolla, CA) labeled with a second fluorophore (Cy3). The fluorescent dyes were purchased from Amersham Pharmacia Biotech (Piscataway, NJ). Hybridizations were carried out using the standard protocol that was described previously [19].

Data Analysis and Clustering

Array images were scanned using an Axon Scanner 4000B (Axon Instruments, Union City, CA), and image analysis was performed using GenePix Pro version 5.0 3.0.6.89 (Axon Instruments). The raw data files were stored in the Stanford Microarray Database [23], and the data that were used for the article are available at: http://microarray-pubs.stanford.edu/tumor-endothelial interaction. Data were expressed as the $\log_2$ ratio of fluorescence intensities of the sample and the reference for each element on the array.

The (Cy5/Cy3) ratio is defined in the Stanford Microarray Database [23] as the normalized ratio of the background-corrected intensities. Spots with aberrant measurements that were due to obvious array artifacts or poor technical quality were manually flagged and removed from further analysis. A filter was applied to omit measurements where the fluorescent signal from the DNA spot was less than 50% above the measured background fluorescence that surrounded the printed DNA spot in either the Cy3 or the Cy5 channel. Genes that did not meet these criteria for at least 80% of the measurements across the experimental samples were excluded from further analysis. Valid data were filtered to exclude elements that did not have at least a three-fold deviation from the mean in at least two samples. Data were evaluated by unsupervised hierarchical clustering [24] and significance analysis of microarray [25] and were displayed using TreeView (http://rana.lbl.gov/EisenSoftware.htm).

GO::Termfinder

GO::TermFinder comprises a set of object-oriented Perl modules for accessing Gene Ontology (GO) information and evaluating and visualizing the collective annotation of a list of genes to GO terms [26]. It can be used to draw conclusions from microarray and other biologic data by calculating the statistical significance of each annotation.

Determination of the Heterotypic Interaction Effect on Gene Expression

To facilitate the identification of heterotypic interaction effects on global gene expression in a mixed coculture experiment, the gene expression data were normalized based on the proportional contribution of each cell type to transcript abundance. Given that the average gene does not change because of the heterotypic interaction and that there are simple additive effects to account for, a linear regression fit was used for normalization. To determine the contribution of each cell type to the combined gene expression pattern in the linear regression model, the expression levels of the monocultures were the predictors, and the expression levels of the coculture were the response.

Specifically, a set of equations $(1 - n)$ was established (one per gene): $e_n^{\text{coculture}} = ((a \times e_n^{\text{monoculture1}}) + ((1 - a) \times e_n^{\text{monoculture2}})) \times I_n$, where e

represented the expression level of the gene, a represented the proportional contribution of mRNA from the respective monoculture, n represented the number of genes measured on the microarray, and I represented the interaction coefficient. We assume that the average gene is not influenced by the heterotypic interaction in the mixed coculture, which is represented as $I = 1$. Because the data set over e_{1_n} is skewed, a linear regression fit was empirically identified based on Gamma errors and identity link as a good model to calculate a. The equation $1 - n$ can then be solved for I_{1_n}, which results in a profile of interaction effects for the genes$_{1_n}$. These interaction effects can be analyzed in much the same way as conventional gene expression measurements.

Human Breast Cancer Data Set

The data set for breast cancer contained 295 tumors that were analyzed on a 25,000-spot oligonucleotide array [27]. In brief, patients were diagnosed and treated at the Netherlands Cancer Institute (NKI) for early-stage breast cancer (stages I and II) between 1984 and 1995. The clinical data were updated in January 2005. The median follow-up for patients who are still alive is 12.3 years.

The "tumor-endothelial cell–induced M-phase cell cycle" gene list consists of 98 genes that are represented by 95 image clones on the cDNA Stanford Array. Clones having the same Unigene locus were removed. The gene sequences were mapped to spots on the NKI array using Unigene build no. 184 (released on June 9, 2005) to give 36 unique spots. To overcome possible overweighting of clones from Unigene clusters that were matched to more than one probe on the NKI array, expression values that were derived from probes that were not matched to the same Unigene cluster were averaged. Expression measurements for each gene were mean centered. The resulting data set was subjected to hierarchical clustering with average linkage clustering [24] and displayed with TreeView (http://rana.lbl.gov/EisenSoftware.htm).

Distant metastasis was analyzed as first event only (distant metastasis-free probability). If a patient developed a local recurrence, an axillary recurrence, a contralateral breast cancer, or a second primary cancer (except for nonmelanoma skin cancer), he/she is censored at that time and the subsequent distant metastases are not analyzed. This is based on the theoretical possibility that the locally recurrent or second primary cancers could be a source for distant metastases. An ipsilateral supraclavicular recurrence was soon followed by a distant metastasis in all but one patient. An ipsilateral supraclavicular recurrence was thus considered to be the first clinical evidence for metastatic disease for this analysis, and patients were not censored at the time of ipsilateral supraclavicular recurrence. Overall survival was analyzed based on death from any cause, and patients were censored at last follow-up. Kaplan-Meier survival curves were compared by the Cox-Mantel log-rank test using Winstat for Microsoft Excel (R Fitch Software, Staufen, Germany).

A data set of gene expression patterns from advanced breast cancers was described by Sorlie et al. [28]. Expression data from 53 image clones that represented the "tumor-endothelial cell–induced M-phase cell cycle" gene list were included in this data set. Genes and samples were organized by hierarchical clustering. Relapse-free and overall survivals were calculated as described above.

Centroid Correlation

The method of calculating the centroid for each patient was previously described by Sorlie et al. [28]. The centroids were profiles that consisted of the average gene expression value for each of the patients. Briefly, the centroids for the genes that represented the "tumor-endothelial cell–induced M-phase cell cycle" signature as well as the other signatures were calculated based on the NKI data set. To test a similarity between the signatures, the correlation between values of different centroids was checked for each patient. The correlation was calculated using the Pearson correlation coefficient with R software [29].

Results

Setup of a Tumor-Endothelial Coculture Model

As a model for investigating the gene expression program in response to epithelial-endothelial interactions in the normal breast and in breast cancer, cells that represented either benign or malignant epithelial cell compartments and cells that represented endothelial cell compartments were examined in an *in vitro* mixed coculture setting. The cells were cocultivated for 48 hours in low-serum medium (0.2% FBS) to allow reciprocal signal exchange with minimal background from the influence of undefined molecular signals that are inherent in fetal bovine serum. We examined the effects of cocultivation for each cell pair in two independent biologic replicates. The gene expression profiles of the cocultures were compared with the expression profiles of the corresponding cells that were kept in monoculture using cDNA microarrays that contained approximately 40,700 elements representing 24,472 unique Unigene clusters (build no. 173, released on July 28, 2004). To establish this experimental approach, we first focused our experiments on the breast cancer cell line, Hs578T, the dermal microvascular endothelial cell, HDMEC, and the coculture of these two cell types. The data passing our data quality filter, a filter for data distribution and a filter selecting genes that are more consistent within replicate samples than between experimental samples, were organized using unsupervised hierarchical clustering of the replicate experiments to provide an overview of the effects on global gene expression (Figure 1*A*). In the coculture, most genes displayed intermediate expression levels, which closely approximated the proportionally weighted average of their expression levels in the two cell types in monoculture. Despite setting up the coculture with equal cell numbers of Hs578T and HDMEC, the gene expression pattern after cocultivation is dominated by the pattern of Hs578T. However, one set of genes showed a consistent increase in transcript abundance in the coculture when compared with either monoculture, which suggested that the induction of these genes was an effect of cocultivation.

Interestingly, 11% of the 44 genes within this gene set are involved in DNA replication as determined by GO terms, and they are *RRM2*, *Cdc45L*, *MCM4*, *KIAA1212*, and *MCM8* [26] (Appendix Figure 1). The frequency of the genes that were involved in this function is significantly enriched ($P = .006$) compared with the background of 1140 genes passing the data quality and data distribution filters as shown in the heat map of Figure 1. Furthermore, several genes are involved in the cell cycle, and they are *CCNA2*, *RAD54B*, *AURKA*, and *CENPN*. *STC1*, which is an extracellular matrix protein found in breast cancer [30] and known to be involved in angiogenesis [31], was also detected in this set of genes. Taken together, these gene sets suggest proliferation of the tumor-endothelial cell coculture.

Because breast cancer is a clinically and molecularly heterogeneous disease, we selected a broad spectrum of different breast cancer cell lines to sample this heterogeneity and explore the effects of a heterotypic

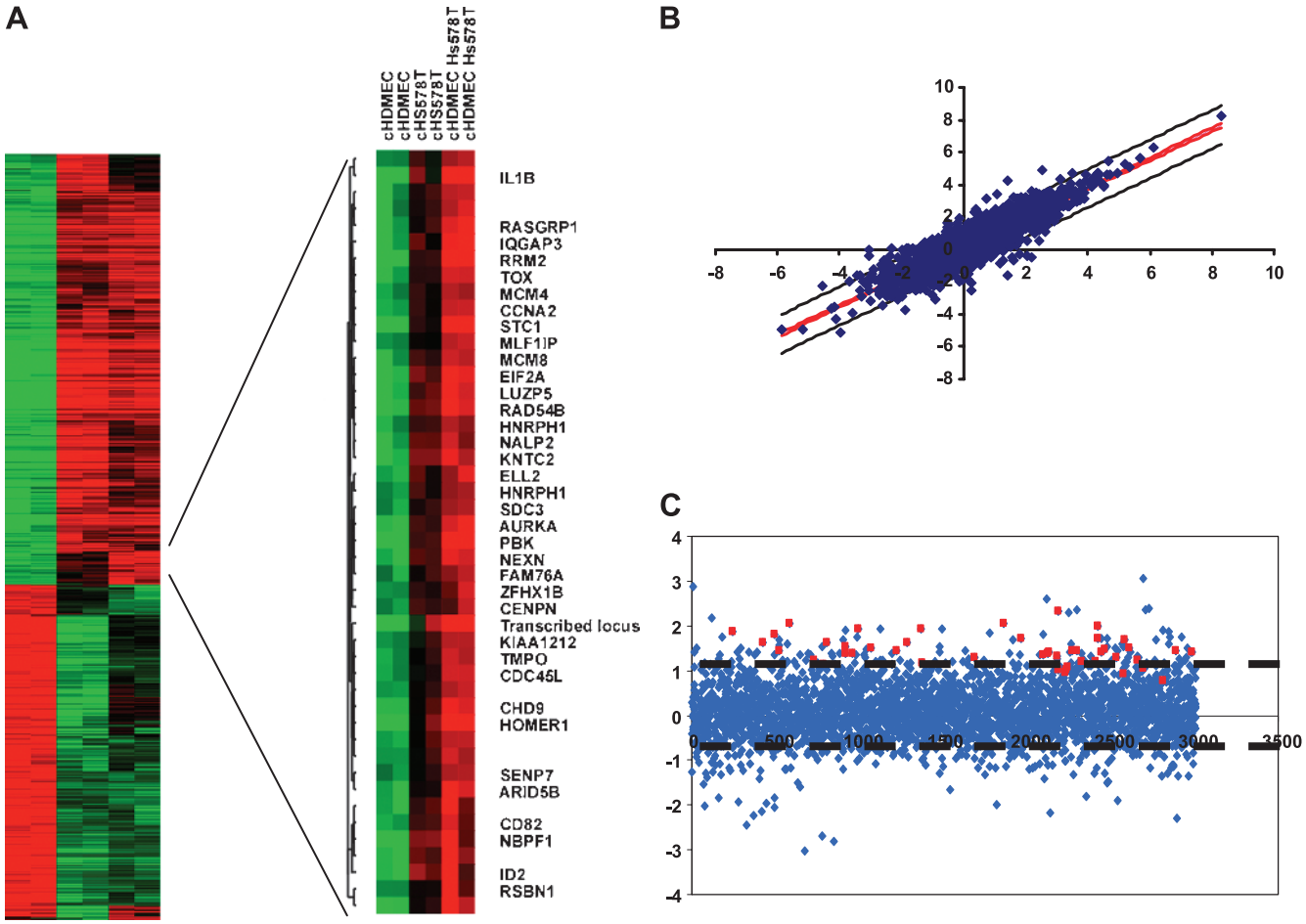


Figure 1. Effect of heterotypic interaction between an endothelial cell and a breast cancer cell line. (A) Biologically independent replicates of the monocultured HDMEC, the breast cancer cell line Hs578T, and the mixed coculture of HDMEC and Hs578T were grown for 48 hours at low serum conditions and characterized by DNA microarray hybridization. Hierarchical clustering of a total of 1140 elements that display a greater than three-fold variance in expression in more than two different experimental samples. Data from individual elements or genes are represented as single rows, and different experiments are shown as columns. Red and green denote the expression levels of the samples. The intensity of the color reflects the magnitude of the deviation from baseline. Unsupervised hierarchical clustering of the experiments grouped the biologic replicates together. Gene expression varied considerably between HDMEC and Hs578T cultures as expected for cells of mesenchymal or epithelial origin, respectively. The coculture profile showed mainly intermediate expression levels. However, the vertical black bar marks a cluster of genes that were induced in all cocultures when compared with both monocultures, which indicated that they were induced by the heterotypic interaction. Zooming in on the genes that were upregulated in coculture revealed that they were specific for proliferation and mitosis. (B) Correlation of the measured coculture gene expression levels and their estimated expression levels based on the proportional contribution of each cell type as determined by a linear regression fit of the monoculture to the coculture data. (C) Fold change of each gene that was associated with coculturing of HDMEC and Hs578T. Genes of the "proliferation and mitosis" cluster are indicated in red. Dashed lines indicate 95% confidence limits.

culture by looking for subtype-specific and shared response patterns. The focus was on epithelial-endothelial cell interactions, which were studied by cocultivating endothelial cells of different origins (HUVEC and HDMEC) in combination with HMECs or six widely used breast cancer cell lines. HDMECs, which are a commercially available dermal microvascular endothelial cell line, were selected to resemble the tumor vasculature endothelial cells of breast cancer as accurately as possible. HUVECs were selected to represent venous cells, despite the fact that the umbilical and breast environments are different. However, the changes that were observed suggest the system works well.

The changes in gene expression that are due to heterotypic interactions were subtle when compared with the large intrinsic variation in expression patterns among the involved cell types, as illustrated in Figure 1*A* for the cell pair of Hs578T and HDMEC. To identify the gene expression changes that resulted from cell-cell interactions, it was necessary to control for the simple additive effects that reflect the proportional contribution of the two cell types to the total population of each gene's transcript in coculture. Elimination of these proportionally weighted additive contributions allowed the isolation of supra-additive interaction effects. The fact that the transcript levels of most genes did not change in response to coculture allowed a linear regression model that was based on the transcription profiles of each monoculture be fitted to the coculture data for normalization. An example of this type of analysis is shown in Figure 1*B*. For each gene, the ratio of the measured transcript level and the level that was estimated by the linear model provides a measure of the heterotypic interaction effect. This is illustrated in Figure 1*C*, which shows the distribution of the gene expression changes of the Hs578T/HDMEC coculture. The genes that

were identified by hierarchical clustering as differentially expressed in coculture when compared with monoculture are highlighted to illustrate the performance of this approach. Interaction effects, which are represented as gene expression changes, are converted to quantitative values that can be analyzed for similarities and disparities over multiple different pairwise interactions between cells with the same tools that are used to analyze conventional gene expression data.

There was an obvious heterogeneity in the responses of different pairs of cells to cocultivation (Figure 2). A striking feature was a cluster of genes that was induced in MDA-MB-231 and Hs578T cocultures with either HUVEC or HDMEC but not in the cocultures with the HMECs. To test for enrichment in a specific functional gene ontology, we applied again the GO::Termfinder tool and found that this set of 104 genes was highly enriched for genes associated with the M-phase of the cell cycle compared with the background of 8140 genes (P = 1.87e – 16). This set of "tumor-endothelial cell–induced M-phase cell cycle" genes included *TOP2A*, *CDC2A*, *CCNB2*, *DHFR*, *CIT*, *TK1*, *CCNA*, *RRM2*, *NUSAP1*, *TPX2*, *CDC25C*, *CENPA*, *CENPE*, *AURKA*, *DTL*, *DICER*, *WHSC1*, *PSIP1*, and *MTPN*. Interestingly, *WHSC1* plays a role in tumorigenicity, adhesion, and clonogenic growth in multiple myeloma [32]; *PSIP1* [33], which is a transcriptional coactivator that is involved in different processes, plays a protective role in stress induced apoptosis; and *MTPN* [34] is a gene that belongs to the myotrophin family.

Differential Induction of Proliferation with Conditioned Medium from Different Tumor Types

As implied by the higher expression of the "tumor-endothelial cell–induced M-phase cell cycle" gene signature, the proliferation rate, as determined by direct cell counting over time, was significantly higher in the coculture of Hs578T and HUVEC than in isolated Hs578T or HUVEC cultures (Figure 3*A*). Also, by flow cytometric cell cycle analysis using propidium iodine staining, we observed a higher proportion of cells leaving the G_1 phase in the coculture than in the monocultures indicative of a higher proportion of cells cycling and passing the M-phase (data not shown).

We speculated that specific factors secreted by the tumors might increase the growth of the endothelial cells or *vice versa*. Toward this aim, we reciprocally incubated one cell type in conditioned medium from the other cell type. In fact, when the conditioned medium from the HUVECs is applied to the tumor cells, the tumor cells show a significantly higher proliferation rate than when kept in autologous medium as measured by the increase in relative cell numbers using the WST-1 proliferation assay (P = 1.3e – 12 [MCF-7], P = 4.6e – 06 [Hs578T], P = 4.3e – 07 [MDA-MB-231], unpaired 2-sided t test; Figure 3*B*). This explains the enhanced proliferation, but it does not explain the differences in proliferation between cocultures of the distinct tumor subtypes. Of note, the conditioned medium from HUVECs induced a higher level of proliferation in MCF-7 cells than in either MDA-MB-231 or Hs578T cells. However, when the conditioned media that are derived from the different tumor types are applied to HUVECs, there is an increase in proliferation on stimulation with conditioned medium from Hs578T cells, which is not seen in the response to the supernatant from MCF-7 cells (Figure 3*C*). This pattern of differential induction is consistent with the differences in the expression of the "tumor-endothelial cell–induced M-phase cell cycle" gene signature, which was observed in the cocultures of Hs578T or MCF-7 in combination with HUVECs.

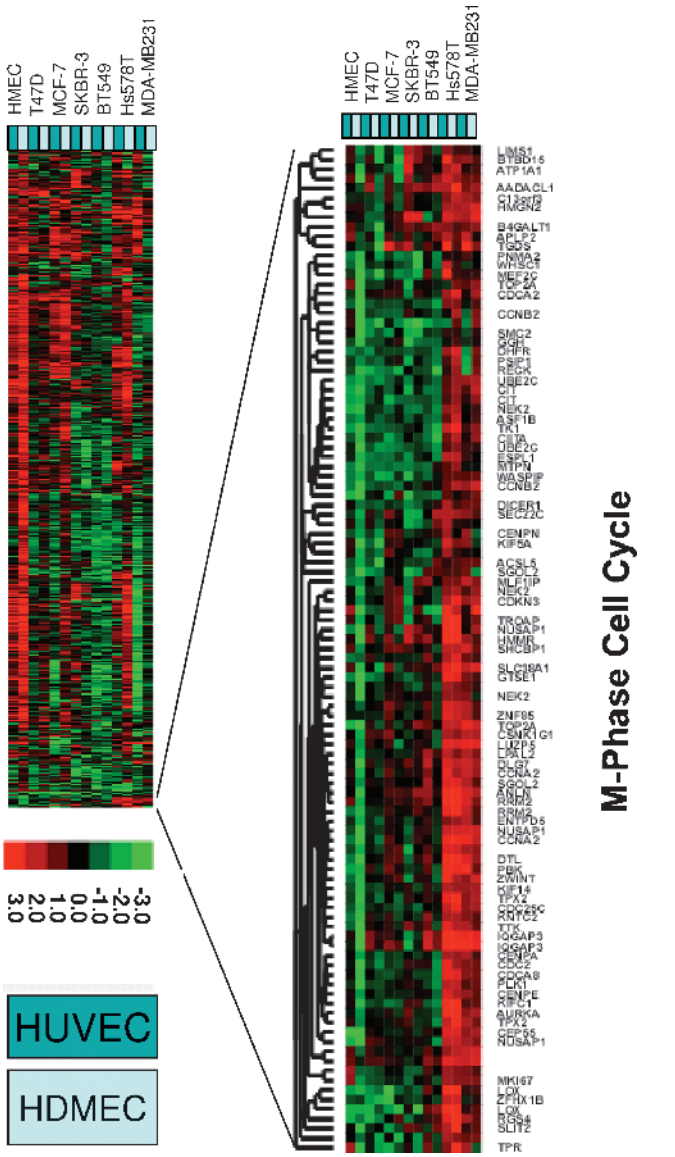


Figure 2. Gene expression changes in multiple cocultures of breast cancer cell lines with endothelial cells. Overview of collapsed data from repeat coculture experiments of eight benign and malignant epithelial cells with two different endothelial cells (HUVEC and HDMEC). The 9 monocultures and the 14 cocultures were analyzed independently in duplicates on 46 HEEBO arrays. Raw data were filtered for technical quality as described in Materials and Methods section, leaving 37,773 spots. A data distribution filter eliminating spots with a $\log_2$-based red-green normalized ratio of less than 1 in at least two arrays removed 10,183 suid(s), leaving 27,590. Genes with missing data in more than 20% of the arrays were removed, leaving 14,565. On the basis of this data set, the calculation of the interaction factors was performed for all 14 cocultures separately as described in Materials and Methods section. The interaction factors were then further analyzed for their distribution, and factors with an SD of less than 0.5 were eliminated, leaving interaction factors for 8140 genes, which are shown as a heat map after unsupervised hierarchical clustering. Red and green denote relative changes in expression that were associated with heterotypic interaction. The magnitude of the relative change is given by color intensity. Zooming in on a cluster of genes that was consistently upregulated in more than two of the cocultures revealed that they were specific for the M-phase of the cell cycle.

Differential Pattern of Secreted Factors in the Tumor Subtypes

We speculated that the induction of the "tumor-endothelial cell–induced M-phase cell cycle" genes in cocultures of Hs578T and MDA-MB-231 is due to cell-cell signaling of factors that are specifically expressed at higher levels in these cells. To systematically identify these molecules, a two-class significance analysis of microarray data was performed [25], with one class formed by Hs578T and MDA-MB-231 and the other by HMEC, HDMEC, HUVEC, BT549, SKBR3, T47D, and MCF-7. Among the genes that were expressed at significantly higher levels in these two cell types were several inducers of angiogenesis, which included *VEGFC*, *FGF12*, *PTN*, and *NF1*. There were also several transcription factors that were coexpressed with these genes, such as *SNAI2* and *ZFHx1B*, which are known to be involved in the epithelial-mesenchymal transition. Interestingly, the PDGFB receptor, which responds to PDGFB that is secreted by endothelial cells, was found in these cells (Figure 4*A*). The full list of genes is given in Appendix Table 1.

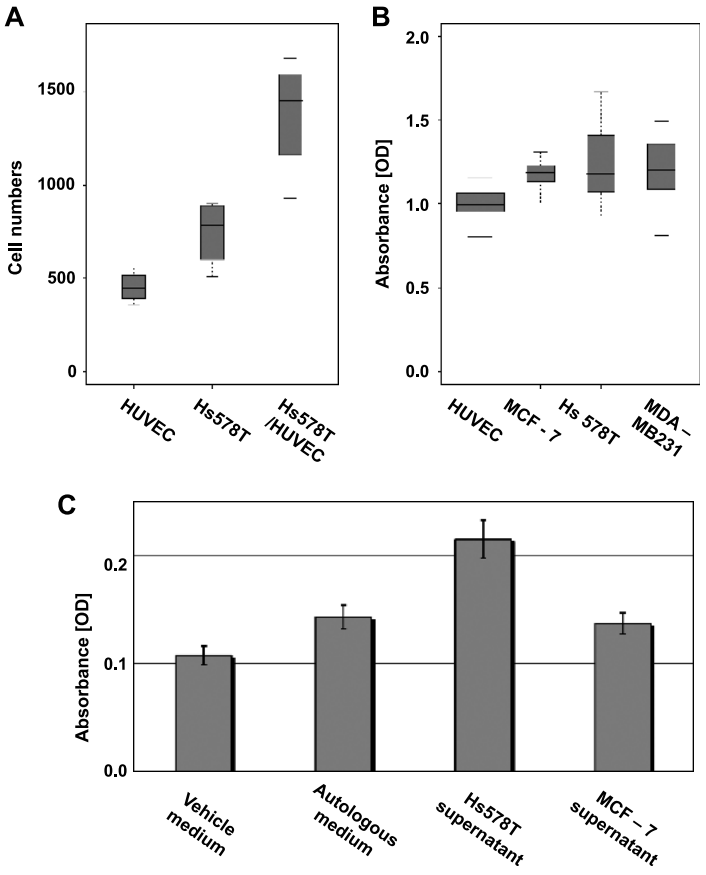


Figure 3. Proliferation of tumor and endothelial cells is due to reciprocal stimulation. (A) Proliferation of HUVECs and Hs578T monocultures and their 1:1 coculture as determined by measuring the increase in cell number by direct cell counting after 36 hours. (B) Box-and-whisker diagrams of relative cell numbers of MCF-7, Hs578T, and MDA-MB-231 after incubation with conditioned medium from HUVECs compared with a normalized negative control of the same cells incubated with autologous medium as measured by the colorimetric cell proliferation assay with the WST-1 compound. (C) Proliferation of HUVECs that was induced by conditioned medium from Hs578T and MCF-7 cells as measured by WST-1. Relative absorbance values of colorful formazan, which has been converted by HUVECs, correspond to relative cell numbers. A single column represents average absorbance values for a minimum of eight independent replicates. *y* Axis error bars correspond to SD. The HUVECs that were treated with the Hs578T supernatant grew significantly faster than the same cells that were treated with the MCF-7 supernatant, the fresh vehicle medium, or the autologous (HUVEC-derived) medium.

Expression of CD44/CD24 in the Diverse Cells

Because breast cancer cells that have the potential to influence the cells within their microenvironment in a way that enhances proliferation might have stem cell characteristics, we speculated that the tumor cells that expressed these angiogenic factors might carry the $CD44^+/CD24^-$ stem cell–like signature, as previously described [35]. Figure 4*B* shows the relative expression of the mRNA of these two antigens in the breast cancer cell lines as measured on the cDNA microarray. CD44 is consistently upregulated in the breast cancer cells that induced the M-phase cell cycle genes in the coculture with endothelial cells, whereas CD24 is consistently downregulated.

This led us to the following working model: Tumor cells first secrete endothelial stimulatory signals, such as VEGF, FGF12, PTN, and NF1 (Figure 4*C*). On stimulation, the endothelial cells start to express PDGFB, which then signals back to the tumor cells through the upregulated PDGFB-R. This feedback loop is facilitated by matching pairs of receptors and ligands, and the coculture starts proliferating.

Blocking Hs578T-Induced Proliferation of HUVEC with Bevacizumab

One of the best described signaling pathways of endothelial cell proliferation is the VEGF pathway [12]. This pathway can be specifically blocked by bevacizumab, which is a monoclonal antibody against VEGF [36]. Whereas Hs578T induces the "tumor-endothelial cell–induced M-phase cell cycle" gene signature in cocultivation with HDMEC, the expression level of this signature in the coculture of MCF-7 and HDMEC represents just the average of the two cells in monoculture (Figure 5*A*). In accordance with these observations, bevacizumab is able to specifically block the stimulatory effect of Hs578T conditioned medium on HUVECs, whereas there was no significant effect on HUVECs that were incubated with conditioned medium that was derived from MCF-7 (Figure 5*B*). Bevacizumab, in contrast to trastuzumab, which served as a negative control, was able to block the conditioned medium that was derived from Hs578T cells and resulted in a reduction in proliferation of HUVECs by approximately 50% to 80%, depending on the antibody concentration (Figure 5*C*). This supports the VEGF pathway as an important factor that enhances tumor-endothelial cell proliferation. However, it is not sufficient alone to explain the full effect. Other factors that were identified by differential expression analysis in our coculture model might also be involved and, therefore, represent valuable targets for therapeutic intervention.

In Vivo *Effects of the "Tumor-Endothelial Cell–Induced M-Phase Cell Cycle" Gene Signature*

We investigated the effect on global gene expression in response to heterotypic cell-cell interaction as a simple, controlled, *ex vivo* model of tumor-endothelial cell interaction. We reasoned that identifying and characterizing gene expression programs that were characteristically induced by the interaction between specific pairs of cells in culture might enable us to recognize and interpret specific features in the expression profiles of human cancers that represent similar interactions between

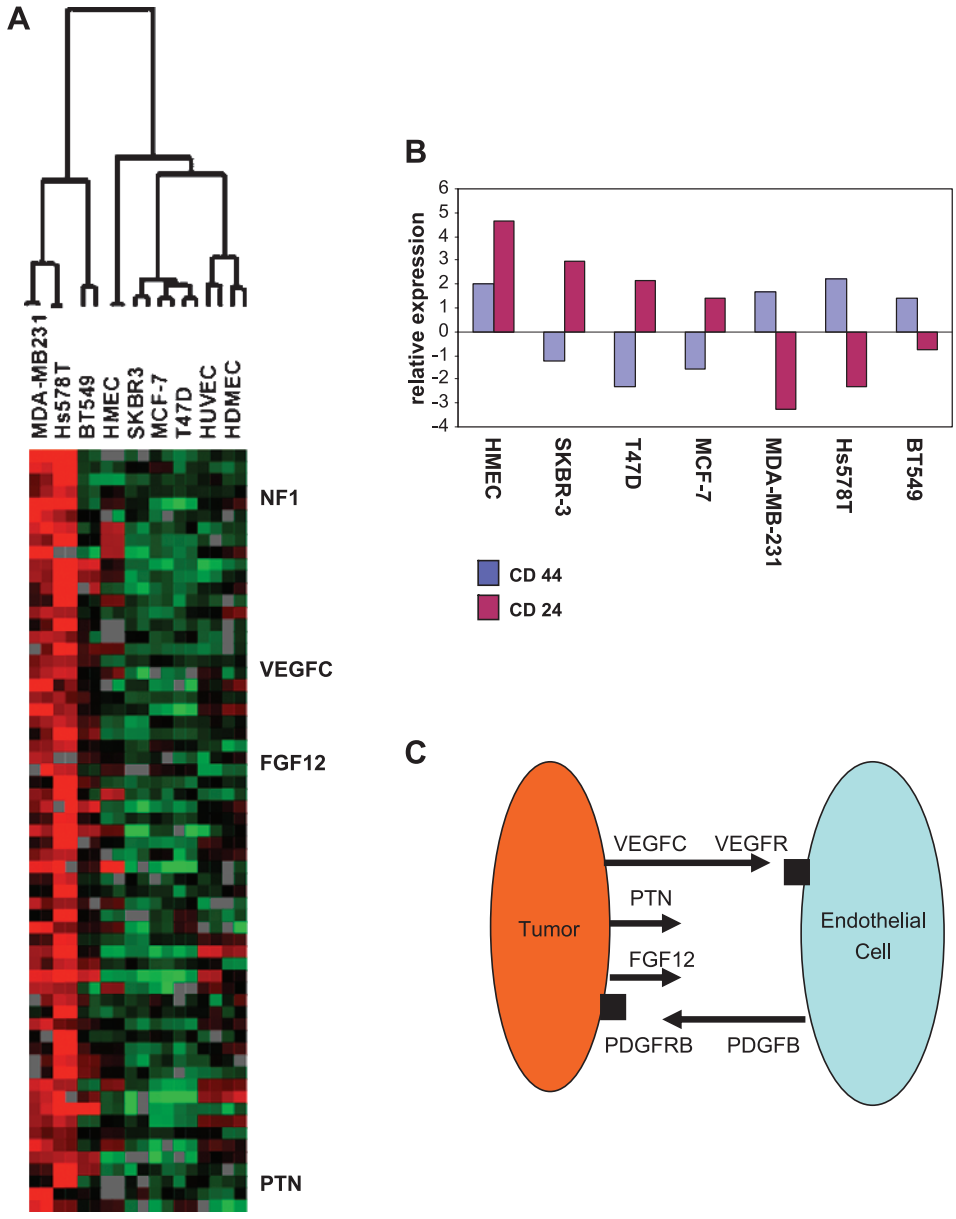


Figure 4. Genes associated with the tumor-endothelial induced M-phase cell cycle gene signature. (A) Significance analysis of microarray data to identify genes that show the largest expression differences between tumor cells that were inducing a proliferation response in coculture with the cells and those that were not inducing proliferation. The expression levels of the top 63 genes are shown on a heat map after unidimensional hierarchical clustering of the genes. (B) Relative expression of CD44 and CD24 in diverse breast cancer cell lines. The average expression of their mRNA over the cell lines corresponds to 0. (C) Working model of reciprocal tumor-endothelial signaling based on the higher expression in the cocultures that induced proliferation.

tumor and endothelial cells *in vivo*. The most consistent response to *ex vivo* cocultivation of breast cancer and endothelial cells was the induction of "tumor-endothelial cell–induced M-phase cell cycle" genes. We, therefore, looked for this response in the expression patterns in the published data from 295 early-stage (stages I and II) breast cancer samples from the NKI [27]. The "tumor-endothelial cell–induced M-phase cell cycle" genes showed a strikingly coherent variation in expression among these cancers, which enabled these cancers to be divided into two groups. One group had a relatively high expression and the other had a relatively low expression of the "tumor-endothelial cell–induced M-phase cell cycle" genes. Clustering the breast carcinomas based only on the expression of the "tumor-endothelial cell–induced M-phase cell cycle" genes separated them into two main clusters, with one cluster having a high-level expression of most of the "tumor-endothelial cell–induced M-phase cell cycle" genes and the other having a lower expression of these genes (Figure 6*A*). The same coordinated behavior and segregation of tumors could be observed in a different set of advanced breast cancer samples [28,37], which suggested that variation in this "tumor-endothelial cell–induced cell M-phase cell cycle" program is a general feature in breast cancer (Appendix Figure 2).

As a first assessment of its potential biologic relevance, distant metastasis-free survival and overall disease-specific survival were compared between the two groups. Early-stage tumors with high expression

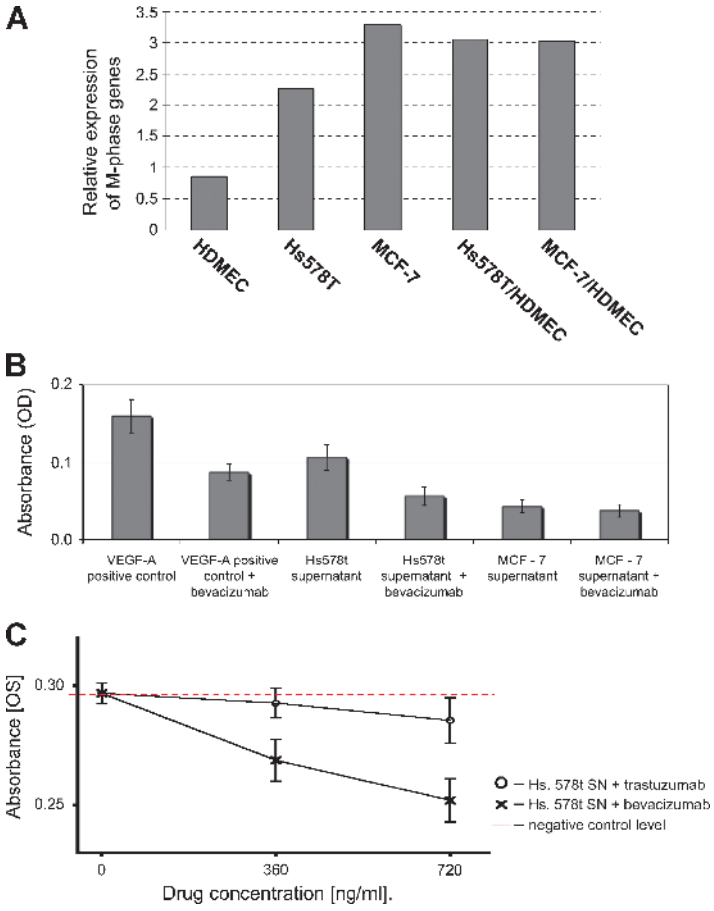


Figure 5. Stimulatory effect of Hs578T conditioned medium on endothelial cells can be partially blocked by bevacizumab. (A) Relative expression of the "tumor-endothelial cell–induced M-phase cell cycle" genes in different monocultures of breast cancer cell lines and endothelial cells and in their respective cocultures. (B) Blocking of the stimulatory effect of Hs578T conditioned medium by bevacizumab. Absolute absorbance values of formazan dye that was converted by HUVECs, which corresponds to cell numbers, are shown in columns, with the *y* axis bars corresponding to SD. Bevacizumab (100 ng/ml) depleted the stimulatory effect of Hs578T cell culture supernatant in a significant manner, whereas it had no effect on the MCF-7 cell culture supernatant. Recombinant VEGF-A (5 ng/ml) and 5% FBS served as positive and negative controls, respectively. (C) Dose-dependent blocking of HUVEC proliferation by bevacizumab and trastuzumab. Absolute absorbance values of formazan dye that was converted by HUVECs are shown. HUVECs that were treated with the Hs578T cell culture supernatant represent the baseline stimulatory effect. Bevacizumab depleted the stimulatory effect of the Hs578T conditioned medium in a significant, dose-dependent manner. Trastuzumab, which is a monoclonal antibody against HER2, did not influence HUVEC stimulation by the Hs578T conditioned medium.

levels (n = 137) of this particular gene set had a significantly lower distant metastasis-free survival (P = 1.8e – 5; 50% at 10 years) and overall survival rate (P = 5e – 9; 52% at 10 years) than tumors with low expression levels (n = 158; metastasis-free survival, 73% at 10 years; overall survival, 84% at 10 years; Figure 6*B*).

The same trend toward unfavorable outcome in patients with cancers that showed high levels of "tumor-endothelial cell–induced M-phase cell cycle" gene transcripts (P = .17) could be seen in the analysis of the data set from advanced-stage breast cancers [28] (data not shown) from Norway/Stanford.

Correlation to Other Prognostic Gene Expression Signatures

The relationship between the "tumor-endothelial cell–induced M-phase cell cycle" gene signature and three previously identified gene expression signatures, which were useful prognosticators in this data set, were also investigated (Figure 7). The first signature is a set of 70 genes [38], which was identified in a supervised analysis of a subset of the NKI early-stage breast cancer data set [27], that could predict freedom from metastasis at 5 years. The second signature was identified *in vitro* by exposing fibroblasts to serum to mimic a wound response, and it has been shown to predict a risk of progression [39]. The "tumor-endothelial cell–induced M-phase cell cycle" gene signature seems to be highly correlated with the poor prognosticator wound (0.605) signature and to be the reciprocal of the good prognosticator 70-gene signature (–0.708), which indicated that our hypothesis that was based

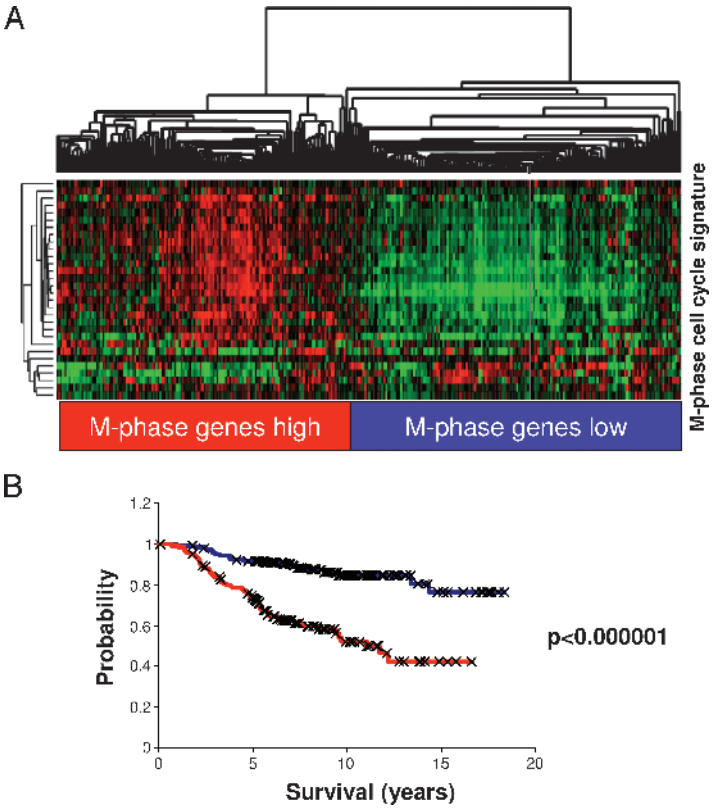


Figure 6. "Tumor-endothelial cell–induced M-phase cell cycle" genes in early-stage breast cancer. (A) The expression values of genes in the "tumor-endothelial cell–induced M-phase cell cycle" gene signature were extracted from a published expression study of 295 early-stage breast cancers from the Netherlands Cancer Institute [35]. Genes and samples are organized by hierarchical clustering. The tumors were segregated into two groups that were defined by high (red) or low (blue) expression levels of the 30 genes matching the M-phase cell cycle gene cluster. (B) Correlation of the "tumor-endothelial cell–induced M-phase cell cycle" gene signature status with distant metastasis-free and overall survival. Kaplan-Meier curves for the clinical outcomes of the indicated tumors that exhibit high (red curve) and low (blue curve) "tumor-endothelial cell–induced M-phase cell cycle" gene signature expression are shown.

on an *in vitro* model is of *in vivo* clinical relevance. Interestingly, the "tumor-endothelial cell–induced M-phase cell cycle" gene signature has even a higher correlation to the 70-gene profile, which was derived as a prognosticator by a top down analysis of this exact data set, than the wound signature. The invasiveness gene signature, which was derived from a comparison of $CD44^+/CD24^-$ stem cell–like cells with normal breast epithelial cells, predicted poor prognosis in breast cancer [40]. This signature, despite having similar prognostic power as the wound signature, did not correlate well with the 70-gene signature, with the wound signature, or with the "tumor-endothelial cell–induced M-phase cell cycle" gene signature (–0.21).

Discussion

The main objective of this study was to examine and characterize the effects of heterotypic cellular interaction to gain insight into the underlying biology of these effects in normal mammary tissue and in breast cancer, with a specific focus on the interaction between epithelial tumor cells and endothelial cells. To isolate specific, direct interactions from more complex interactions that involve multiple cell types in a whole tissue or organism, we used a simple *ex vivo* coculture model. Because some important heterotypic interactions might require direct cell-cell contact, we focused on a coculture model where the two cell types were mixed [22]. In this report, we describe the systematic genomic analysis

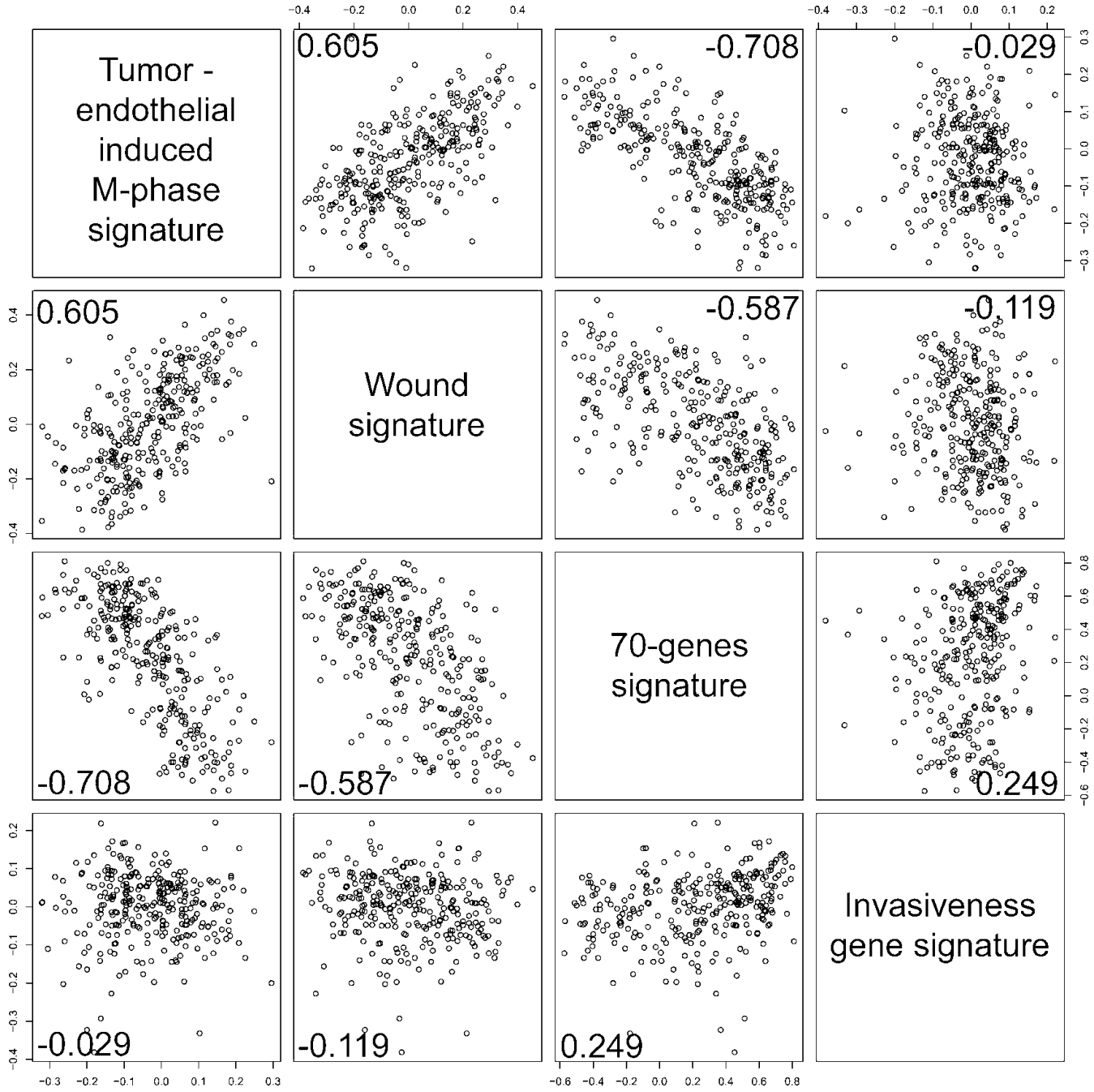


Figure 7. Correlation to other prognostic gene signatures in early-stage breast cancer. Correlation of the 70-gene signature [50], the wound signature [49], the invasiveness gene signature [40], and the "tumor-endothelial cell–induced M-phase cell cycle" signature score in the NKI data set. Pairwise scatter plot matrix of the three gene signatures. Pearson correlations are shown in the lower part of each plot.

of this simple *in vitro* system that simulates direct and indirect interactions between benign and malignant epithelial cells and endothelial cells in normal breast tissue and in breast cancer.

Common Gene Set Induced by Tumor-Endothelial Coculture

As expected, based on our experience with the tumor-fibroblast interaction [22], the picture of heterotypic interaction effects from the survey of diverse tumor cells with two different types of endothelial cells is complex and reflects the different abilities of normal and malignant cells to send and respond to extrinsic signals. Our data show that the effects of the tumor-endothelial cell interaction differ between different breast cancer cell lines that each represents a different breast cancer subtype. A prominent theme in our coculture experiments was a set of genes that were characteristic of the mitotic phase of the cell cycle, which we called the "tumor-endothelial cell–induced M-phase cell cycle" gene expression signature. This observation on the gene expression level is consistent with the phenotypic features that show a higher proliferation rate in the respective cocultures. In our setting, the mere coexistence of two cell types of different origin, such as the Hs578T breast cancer cells and endothelial cells, seems to be sufficient to induce proliferation. Cooperate induction between cells of different lineages is well known from developmental biology, where stem cells cooperate with their environmental cells to form the stem cell niche. The stem cell concept has also been introduced in cancer [41]. Cells with stem cell characteristics have been prospectively isolated from breast cancer using CD44 and CD24 as markers [35]. From these cells, which are characterized as $CD44^+/CD24^-$, a prognostic gene expression signature has been determined, which has been called the invasiveness signature [40]. This signature has been associated with an unfavorable prognosis in breast cancer. However, it was unclear how that small fraction of stem cells could lead to such a prominent gene expression pattern in the genomic profile of breast cancer, and it was unexpected that a signature that was derived from stem cells, which are known for their low proliferative potential, was linked to the highly proliferative, poor prognostic tumors.

Hs578T and MDA-MB-231 both exhibit the $CD44^+/CD24^-$ phenotype. Whereas in monoculture under low serum conditions, they proliferated slowly in a manner that was similar to stem cells [42,43], in coculture with endothelial cells, they had a high proliferation rate along with associated changes in gene expression. This induction of proliferation markers is absent in the cocultures with $CD44^-/CD24^+$ cells. Therefore, cocultivation of endothelial cells with breast cancer cells links the $CD44^+/CD24^-$ signature with high proliferation, which is then associated with the overexpression of the "tumor-endothelial cell–induced M-phase cell cycle" gene signature and predicts a poor prognosis in breast cancer. We are well aware that the $CD44^+/CD24^-$ signature is preferentially associated with basal type carcinomas, in contrast to the $CD44^-/CD24^+$ signature, which is associated with the luminal-type breast cancer cells [44]. In this context, the cocultivation with endothelial cells explains the worse prognosis of basal type breast cancer despite the lower proliferation rate of the isolated basal-type breast cancer cells in monoculture. This exemplifies that this simple *in vitro* coculture model is more similar to the *in vivo* situation than the monoculture. The stem cell–like cells with highly potent tumor-initiating properties, as they were described by Al-Hajj et al. [35], clearly exhibited the $CD44^+/CD24^-$ signature but were also characterized by additional markers. It has been shown that $CD44^+/CD24^-$–expressing cell lines contain these tumor-initiating cells [42,45], but, of note, we did neither specifically focus on nor explicitly select for these stem cell–like cells.

Differences Are Due to Differential Endothelial Cell Proliferation

A comparison of the gene expression program between the $CD44^+/CD24^-$ tumor cells that upregulate the "tumor-endothelial cell–induced M-phase cell cycle" genes and the $CD44^-/CD24^+$ tumor cells that do not reveals multiple genes that seem to be involved in angiogenesis. This gene subset includes *VEGF*, *FGF*, and other endothelial stimulatory factors. Many genes in this cluster are functionally less well described, although they might be involved in tumor cell proliferation or in the induction of tumor angiogenesis. To demonstrate their functional impact, we blocked VEGF with bevacizumab, which is a specific antibody against VEGF. A good part of the stimulatory effect of Hs578T conditioned medium on HUVECs was abrogated with bevacizumab. This demonstrates that these genes play a considerable functional role and are not just surrogate markers that are associated with faster tumor growth. Whereas VEGF partially blocks the effects of the tumor cell supernatant, other factors might also contribute to endothelial proliferation and thereby represent possible additional new therapeutic targets. Most notably, these additional factors could be responsible for drug resistance to anti-VEGF monotherapy if tumors start using these alternative angiogenic factors to grow blood vessels.

With consideration to the cancer stem cell concept, the stem cell niche has been proposed to play an important role in carcinogenesis and progression, although it remains largely uncharacterized at the cellular and molecular levels. It is possible to speculate that endothelial cells participate in the breast cancer stem cell niche. In a mouse model, $VEGFR^+$ bone marrow–derived cells have been shown to form a stem cell niche for cancer cells, and such cells have been identified in human breast cancer biopsies [46]. In brain cancer, stem cells were demonstrated to live in a vascular niche that secretes factors that promote their long-term growth and self-renewal [47]. The factors that we have described to enhance the proliferation of Hs578T, a $CD44^+/CD24^-$–expressing cell line, which contains at least in part stem cell–like cells [42] and endothelial cells, might also be involved in a breast cancer stem cell niche. We are well aware that the endothelial cells of the vascular system are diverse [48] and that the tumor vasculature consists of specifically altered endothelial cells. The endothelial cells that we selected, which were the HUVECs and HDMECs, do not represent the autologous tumor vasculature endothelial cells of breast carcinomas. We can, therefore, not exclude the possibility that endothelial cells that are isolated from within a tumor might show additional specific interaction effects. Nevertheless, it would be surprising if carcinoma-associated endothelial cells failed to show the strong effects that we consistently observed in cocultures with HUVECs and HDMECs. Because these experiments might be insufficient to detect subtle differences between cocultures that involve different types of tumor-endothelial cells, a more specific selection of primary tumor-associated endothelial cells would be needed in the optimal case in cocultivation with primary tumor cells.

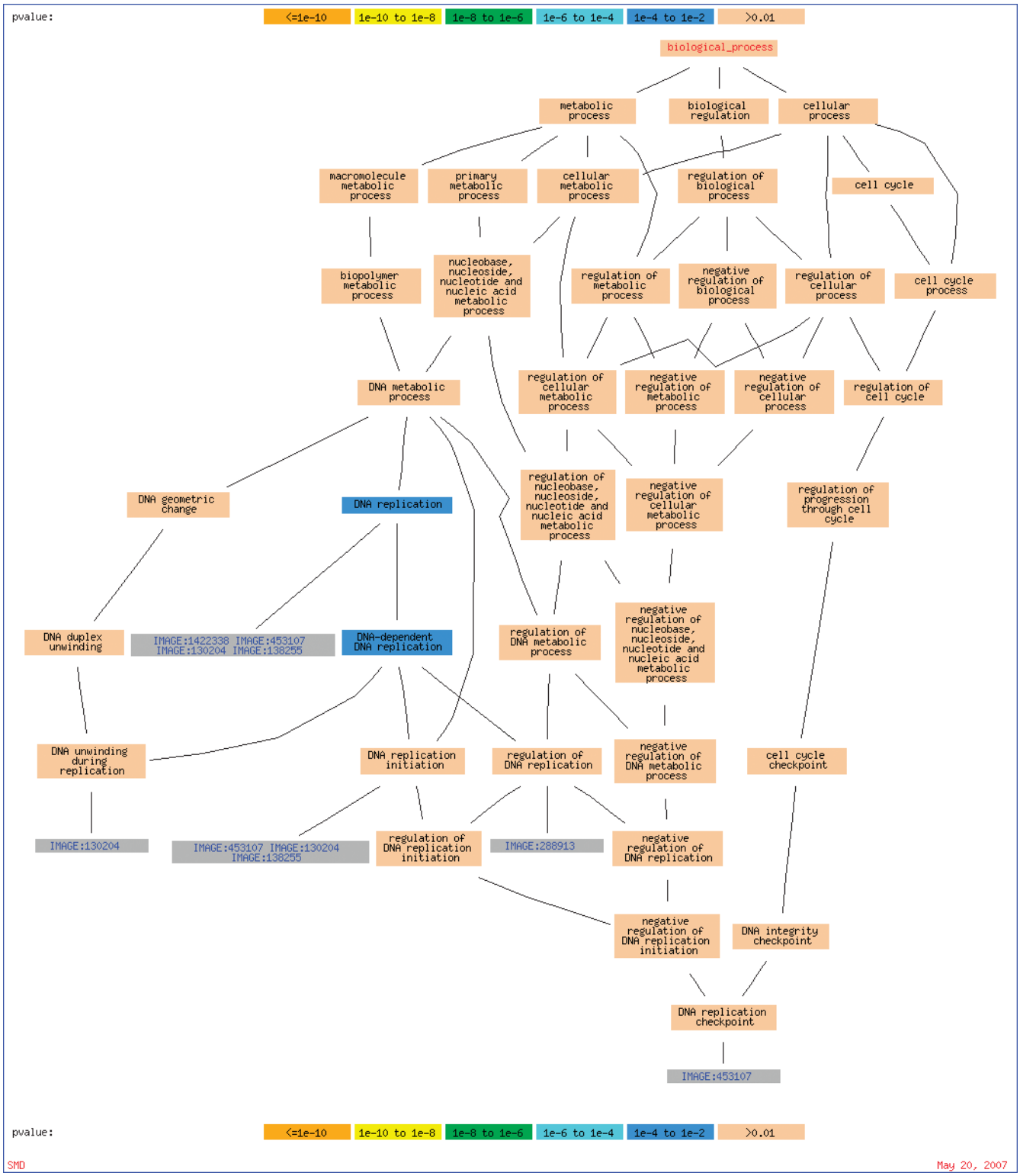
Predictive Marker for Antiangiogenic Drugs

We have shown that the "tumor-endothelial cell–induced M-phase cell cycle" signature is a strong prognostic signature that strongly correlates with other prognostic signatures such as the wound signature [49] and the 70-gene signature [27]. A subset of the genes that are found in our signature overlaps with a subset of the genes that are found in the "wound signature," which was characterized by the stimulation of

fibroblasts with serum, resembling the response of fibroblasts to wounding. Our situation also resembles the situation of wounding, where new vessels have to spread to the wound area to nourish the new wound fibroblasts. The 70-gene signature is currently studied in prospective clinical trials to spare some women chemotherapy without risking their chance of cure [50]. For clinical decision making, it is important to define biomarkers that could serve as predictive factors about the benefit of certain therapies, especially in the case of new therapeutic options, such as antiangiogenic drugs. So far, there is no valid biomarker for this purpose. Although there is evidence from preclinical studies, we have to be cautious when extrapolating preclinical data to the human disease. Preclinical experiments are often designed to look at the influence of well-defined biologic phenomena on drug efficacy. By contrast, human disease is more complex and might require multiple markers to accurately predict efficacy. For this purpose, a set of *in vitro*–designed biomarkers, which are already independently tested for their prognostic power, such as our "tumor-endothelial cell–induced M-phase cell cycle" gene signature or the associated set of angiogenesis-inducing factors, could serve as potential predictors that are worth further evaluation. Our *in vitro* model, which involves the mere coculture of two cell types, seems to allow the identification of a strongly prognostic gene signature. Whereas the genes from the 70-gene signature have not yet been associated with a single functional ontology, our signature is linked to a mechanistic *in vitro* model that allows for further experimentation.

In our previous work, we have modeled the interaction of tumor cells with fibroblasts, which allowed us to define an interferon response gene set [22]. In this work, we have characterized a tumor-endothelial cell interaction. It would be interesting to see how different carcinoma-associated fibroblasts and endothelial cells interact and how the addition of tumor cells would reprogram the system. Our coculture technique may allow us to further explore these more complex interactions among the multiple molecules that operate in these cells to orchestrate the process of cancer progression and metastasis. Our experience suggests that *in vitro* modeling of specific processes and features of the tumor microenvironment can provide a valuable interpretive framework for the analysis of associated gene expression patterns in more heterogeneous *in vivo* samples and the identification of the effects of heterotypic cellular interactions.

Appendix

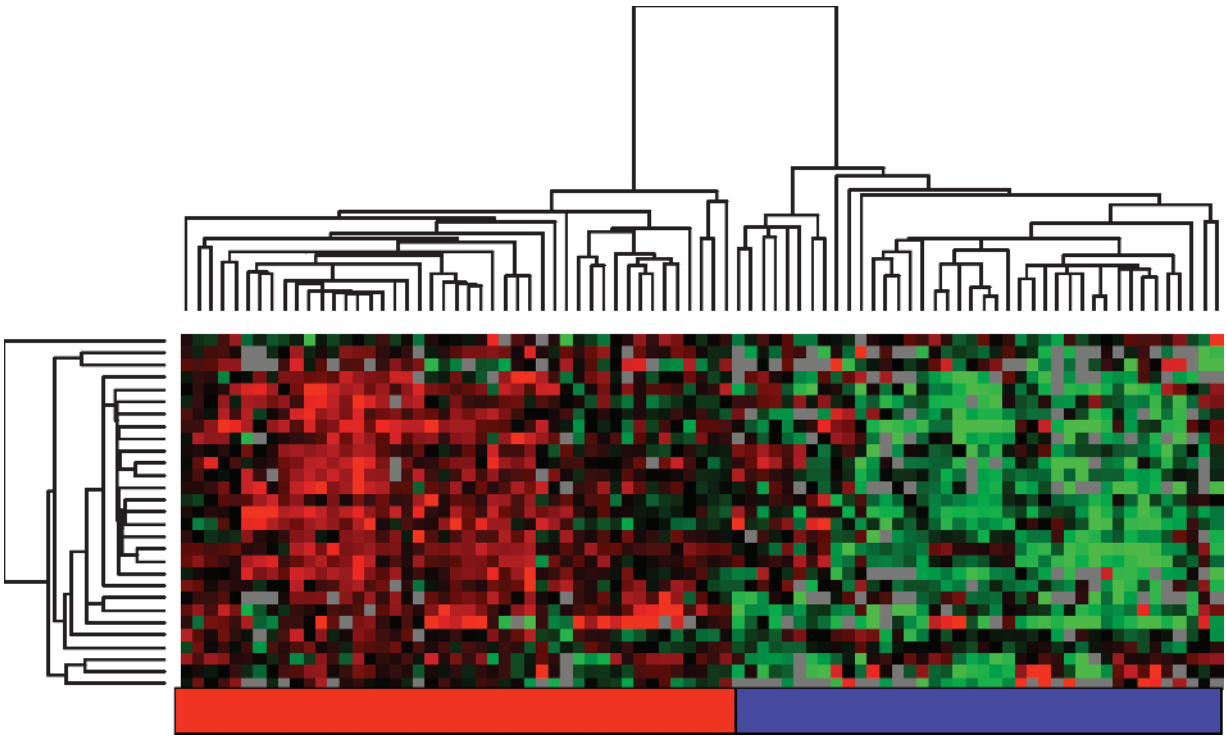


Appendix Figure 1. GO terms. Graphical visualization of the output from GO::Termfinder for process ontology: GOgraph layout that includes the significant GO nodes annotated by DNA replication, derived from 44 clones from a background of 1140. The color of the nodes is an indication of their Bonferroni corrected *P* value (orange ≤ 1e − 10; yellow = 1e − 10 to 1e−8; green = 1e−8 to 1e−6; cyan = 1e−6 to 1e−4; blue = 1e−4 to 1e−2; tan > 0.01).

Appendix Table 1. List of Genes Expressed in Breast Cancer Cell Lines Inducing Proliferation in Coculture.

CLID	UniGene Cluster ID || Symbol || Gene Name
IMAGE:754582	|| Hs.567266 || *NF1* || Neurofibromin 1 (neurofibromatosis, von Recklinghausen disease, Watson disease)
IMAGE:280699	|| Hs.563491 || *EPDR1* || Ependymin related protein 1 (zebrafish)
IMAGE:138991	|| Hs.233240 || *COL6A3* || Collagen, type VI, alpha 3
IMAGE:878836	|| || || AA670429 || || || IMAGE:878836 || 110839
IMAGE:45138	|| Hs.435215 || *VEGFC* || Vascular endothelial growth factor C
IMAGE:1031532	|| Hs.226780 || *OSTM1* || Osteopetrosis associated transmembrane protein 1
IMAGE:460002	|| Hs.269027 || *GALNT5* || UDP-*N*-acetyl-alpha-D-galactosamine:polypeptide *N*-acetylgalactosaminyltransferase 5 (GalNAc-T5)
IMAGE:204737	|| Hs.360174 || *SNAI2* || Snail homolog 2 (*Drosophila*)
IMAGE:1913366	|| Hs.128013 || *PRSS3* || Protease, serine, 3 (mesotrypsin)
IMAGE:898218	|| Hs.450230 || *IGFBP3* || Insulin-like growth factor binding protein 3
IMAGE:489519	|| Hs.297324 || *TIMP3* || TIMP metallopeptidase inhibitor 3 (Sorsby fundus dystrophy, pseudoinflammatory)
IMAGE:489089	|| Hs.474053 || *COL6A1* || Collagen, type VI, alpha 1
IMAGE:488404	|| Hs.27621 || || Clone TUA8 Cri-du-chat region mRNA
IMAGE:753587	|| Hs.167741 || *BTN3A3* || Butyrophilin, subfamily 3, member A3
IMAGE:857640	|| Hs.420269 || *COL6A2* || Collagen, type VI, alpha 2
IMAGE:1030805	|| Hs.458623 || || Transcribed locus, moderately similar to XP_509196.1 PREDICTED: similar to FTO protein [Pan troglodytes]
IMAGE:868212	|| Hs.369397 || *TGFBI* || Transforming growth factor, beta-induced, 68 kDa
IMAGE:2545711	|| Hs.592971 || || Transcribed locus
IMAGE:1470128	|| Hs.411391 || LOC399959 || Hypothetical gene supported by BX647608
IMAGE:121981	|| Hs.419240 || *SLC2A14* || Solute carrier family 2 (facilitated glucose transporter), member 14
IMAGE:1893136	|| Hs.596112 || || Transcribed locus || AI278518 || || || IMAGE:1893136 || 313617
IMAGE:753467	|| Hs.419240 || *SLC2A14* || Solute carrier family 2 (facilitated glucose transporter), member 14
IMAGE:789147	|| Hs.511915 || *ENO2* || Enolase 2 (gamma, neuronal)
IMAGE:299539	|| Hs.584758 || *FGF12* || Fibroblast growth factor 12
IMAGE:471196	|| Hs.111577 || *ITM2C* || Integral membrane protein 2C
IMAGE:504761	|| Hs.374774 || *ANKRD29* || Ankyrin repeat domain 29
IMAGE:811000	|| Hs.514535 || *LGALS3BP* || Lectin, galactoside-binding, soluble, 3 binding protein
IMAGE:153646	|| Hs.172928 || *COL1A1* || Collagen, type I, alpha 1
IMAGE:897768	|| Hs.476218 || *COL7A1* || Collagen, type VII, alpha 1 (epidermolysis bullosa, dystrophic, dominant and recessive)
IMAGE:263716	|| Hs.474053 || *COL6A1* || Collagen, type VI, alpha 1
IMAGE:269425	|| Hs.34871 || *ZFHX1B* || Zinc finger homeobox 1b
IMAGE:754106	|| Hs.297324 || *TIMP3* || TIMP metallopeptidase inhibitor 3 (Sorsby fundus dystrophy, pseudoinflammatory)
IMAGE:809719	|| Hs.477128 || *CCDC80* || Coiled-coil domain containing 80
IMAGE:1558164	|| Hs.89901 || *PDE4A* || Phosphodiesterase 4A, cAMP-specific (phosphodiesterase E2 dunce homolog, *Drosophila*)
IMAGE:293339	|| Hs.360174 || *SNAI2* || Snail homolog 2 (*Drosophila*)
IMAGE:291290	|| Hs.558402 || *SSX4* || Synovial sarcoma, X breakpoint 4
IMAGE:609332	|| Hs.409662 || *COL14A1* || Collagen, type XIV, alpha 1 (undulin)
IMAGE:239256	|| Hs.173859 || *FZD7* || Frizzled homolog 7 (*Drosophila*)
IMAGE:210717	|| Hs.1501 || *SDC2* || Syndecan 2 (heparan sulfate proteoglycan 1, cell surface-associated, fibroglycan)
IMAGE:415122	|| Hs.48384 || *HS3ST3B1* || Heparan sulfate (glucosamine) 3-*O*-sulfotransferase 3B1
IMAGE:625234	|| Hs.642759 || *KDELR3* || KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3
IMAGE:843222	|| Hs.210283 || *COL5A1* || Collagen, type V, alpha 1
IMAGE:234736	|| Hs.514746 || *GATA6* || GATA binding protein 6
IMAGE:80643	|| Hs.482730 || *EDIL3* || EGF-like repeats and discoidin I-like domains 3
IMAGE:502689	|| Hs.632387 || *NEXN* || Nexilin (F actin binding protein)
IMAGE:109424	|| Hs.103110 || *PPARA* || Peroxisome proliferative activated receptor, alpha
IMAGE:813823	|| Hs.406475 || *LUM* || Lumican
IMAGE:110503	|| Hs.480712 || *LARP2* || La ribonucleoprotein domain family, member 2
IMAGE:1475595	|| Hs.75431 || *ALPL* || Alkaline phosphatase, liver/bone/kidney
IMAGE:742125	|| Hs.65436 || *LOXL1* || Lysyl oxidase-like 1
IMAGE:510729	|| Hs.437040 || *PTPN21* || Protein tyrosine phosphatase, non-receptor type 21
IMAGE:769686	|| Hs.643513 || *THY1* || Thy-1 cell surface antigen
IMAGE:144916	|| Hs.62661 || *GBP1* || Guanylate binding protein 1, interferon-inducible, 67 kDa
IMAGE:345849	|| Hs.102267 || *LOX* || Lysyl oxidase
IMAGE:502664	|| Hs.35861 || *TMEM158* || Ras-induced senescence 1
IMAGE:786265	|| Hs.501928 || *MICAL2* || Microtubule associated monoxygenase, calponin and LIM domain containing 2
IMAGE:590759	|| Hs.105269 || *SC4MOL* || Sterol-C4-methyl oxidase-like
IMAGE:854678	|| Hs.567598 || *LBH* || Hypothetical protein DKFZp566J091
IMAGE:344272	|| Hs.9999 || *EMP3* || Epithelial membrane protein 3
IMAGE:361974	|| Hs.371249 || *PTN* || Pleiotrophin (heparin binding growth factor 8, neurite growth-promoting factor 1)
IMAGE:1593317	|| Hs.509067 || *PDGFRB* || Platelet-derived growth factor receptor, beta polypeptide
IMAGE:756372	|| Hs.438823 || *KCNH2* || Potassium voltage-gated channel, subfamily H (eag-related), member 2
IMAGE:415134	|| Hs.632256 || *STAT5B* || Signal transducer and activator of transcription 5B

Genes that are differentially expressed between breast cancer cell lines inducing the "tumor-endothelial cell–induced M-phase cell cycle" gene signature contain multiple endothelial growth factors.



Appendix Figure 2. "Tumor-endothelial cell–induced M-phase cell cycle" genes in advanced-stage breast cancer. The expression values of genes in the "tumor-endothelial cell–induced M-phase cell cycle" gene signature were extracted from a published expression study of advanced-stage breast cancers from Norway/Stanford [33]. Genes and samples are organized by hierarchical clustering. The tumors segregated into two groups defined by high (red) or low (blue) expression levels of 29 genes matching the "tumor-endothelial cell–induced M-phase cell cycle" gene signature.

Acknowledgments

The authors thank Michael Fero, Elena Seraia, and the staff of the Stanford Functional Genomics Facility for supplying the human cDNA microarrays that were used for this study. The authors thank Kathy Ball, Janos Demeter, and the staff from the Stanford Microarray Database for support with the microarray data storage and analysis tools. M.B. is a fellow of the Schoenmakers-Müller Foundation, Basel, Switzerland.

References

[1] Folkman J (1995). Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nat Med* **1**, 27–31.

[2] Folkman J and Shing Y (1992). Angiogenesis. *J Biol Chem* **267**, 10931–10934.

[3] Paweletz N and Knierim M (1989). Tumor-related angiogenesis. *Crit Rev Oncol Hematol* **9**, 197–242.

[4] Fidler IJ and Ellis LM (1994). The implications of angiogenesis for the biology and therapy of cancer metastasis. *Cell* **79**, 185–188.

[5] Kerbel R and Folkman J (2002). Clinical translation of angiogenesis inhibitors. *Nat Rev Cancer* **2**, 727–739.

[6] Yang JC, Haworth L, Sherry RM, Hwu P, Schwartzentruber DJ, Topalian SL, Steinberg SM, Chen HX, and Rosenberg SA (2003). A randomized trial of bevacizumab, an anti–vascular endothelial growth factor antibody, for metastatic renal cancer. *N Engl J Med* **349**, 427–434.

[7] Hurwitz H, Fehrenbacher L, Novotny W, Cartwright T, Hainsworth J, Heim W, Berlin J, Baron A, Griffing S, Holmgren E, et al. (2004). Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med* **350**, 2335–2342.

[8] Sandler A, Gray R, Perry MC, Brahmer J, Schiller JH, Dowlati A, Lilenbaum R, and Johnson D (2006). Paclitaxel-carboplatin alone or with bevacizumab for non–small-cell lung cancer. *N Engl J Med* **355**, 2542–2550.

[9] Escudier B, Pluzanska A, Koralewski P, Ravaud A, Bracarda S, Szczylik C, Chevreau C, Filipek M, Melichar B, Bajetta E, et al. (2007). Bevacizumab plus interferon α-2a for treatment of metastatic renal cell carcinoma: a randomised, double-blind phase III trial. *Lancet* **370**, 2103–2111.

[10] Miller K, Wang M, Gralow J, Dickler M, Cobleigh M, Perez EA, Shenkier T, Cella D, and Davidson NE (2007). Paclitaxel plus bevacizumab *versus* paclitaxel alone for metastatic breast cancer. *N Engl J Med* **357**, 2666–2676.

[11] Rak J, Filmus J, and Kerbel RS (1996). Reciprocal paracrine interactions between tumour cells and endothelial cells: the "angiogenesis progression" hypothesis. *Eur J Cancer* **32A**, 2438–2450.

[12] Ferrara N, Houck K, Jakeman L, and Leung DW (1992). Molecular and biological properties of the vascular endothelial growth factor family of proteins. *Endocr Rev* **13**, 18–32.

[13] O'Reilly MS, Boehm T, Shing Y, Fukai N, Vasios G, Lane WS, Flynn E, Birkhead JR, Olsen BR, and Folkman J (1997). Endostatin: an endogenous inhibitor of angiogenesis and tumor growth. *Cell* **88**, 277–285.

[14] Compagni A, Wilgenbus P, Impagnatiello MA, Cotten M, and Christofori G (2000). Fibroblast growth factors are required for efficient tumor angiogenesis. *Cancer Res* **60**, 7163–7169.

[15] Ergun S, Kilik N, Ziegeler G, Hansen A, Nollau P, Gotze J, Wurmbach JH, Horst A, Weil J, Fernando M, et al. (2000). CEA-related cell adhesion molecule 1: a potent angiogenic factor and a major effector of vascular endothelial growth factor. *Mol Cell* **5**, 311–320.

[16] Akhta N, Carlso S, Pesarini A, Ambulos N, and Passaniti A (2001). Extracellular matrix–derived angiogenic factor(s) inhibit endothelial cell proliferation, enhance differentiation, and stimulate angiogenesis *in vivo*. *Endothelium* **8**, 221–234.

[17] Bissell MJ and Radisky D (2001). Putting tumours in context. *Nat Rev Cancer* **1**, 46–54.

[18] Bissell MJ, Radisky DC, Rizki A, Weaver VM, and Petersen OW (2002). The organizing principle: microenvironmental influences in the normal and malignant breast. *Differentiation* **70**, 537–546.

[19] Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, et al. (2000). Molecular portraits of human breast tumours. *Nature* **406**, 747–752.

[20] Teicher BA (1994). Hypoxia and drug resistance. *Cancer Metastasis Rev* **13**, 139–168.
[21] Li L and Neaves WB (2006). Normal stem cells and cancer stem cells: the niche matters. *Cancer Res* **66**, 4553–4557.
[22] Buess M, Nuyten DS, Hastie T, Nielsen T, Pesich R, and Brown PO (2007). Characterization of heterotypic interaction effects *in vitro* to deconvolute global gene expression profiles in cancer. *Genome Biol* **8**, R191.
[23] Demeter J, Beauheim C, Gollub J, Hernandez-Boussard T, Jin H, Maier D, Matese JC, Nitzberg M, Wymore F, Zachariah ZK, et al. (2007). The Stanford Microarray Database: implementation of new analysis tools and open source release of software. *Nucleic Acids Res* **35**, D766–D770.
[24] Eisen MB, Spellman PT, Brown PO, and Botstein D (1998). Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* **95**, 14863–14868.
[25] Tusher VG, Tibshirani R, and Chu G (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* **98**, 5116–5121.
[26] Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, and Sherlock G (2004). GO::TermFinder-open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* **20**, 3710–3715.
[27] van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, et al. (2002). A gene-expression signature as a predictor of survival in breast cancer [see comment]. *New Engl J Med* **347**, 1999–2009.
[28] Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, et al. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* **100**, 8418–8423.
[29] R Development Core Team (2005). *R: A language and environment for statistical computing, reference index version 2.2.1*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Availbale at http://www.R-project.org.
[30] Wascher RA, Huynh KT, Giuliano AE, Hansen NM, Singer FR, Elashoff D, and Hoon DS (2003). Stanniocalcin-1: a novel molecular blood and bone marrow marker for human breast cancer. *Clin Cancer Res* **9**, 1427–1435.
[31] Gerritsen ME, Soriano R, Yang S, Ingle G, Zlot C, Toy K, Winer J, Draksharapu A, Peale F, Wu TD, et al. (2002). *In silico* data filtering to identify new angiogenesis targets from a large *in vitro* gene profiling data set. *Physiol Genomics* **10**, 13–20.
[32] Lauring J, Abukhdeir AM, Konishi H, Garay JP, Gustin JP, Wang Q, Arceci RJ, Matsui W, and Park BH (2008). The multiple myeloma associated *MMSET* gene contributes to cellular adhesion, clonogenic growth, and tumorigenicity. *Blood* **111**, 856–864.
[33] Ge H, Si Y, and Roeder RG (1998). Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. *EMBO J* **17**, 6723–6729.
[34] Mukherjee DP, McTiernan CF, and Sen S (1993). Myotrophin induces early response genes and enhances cardiac gene expression. *Hypertension* **21**, 142–148.
[35] Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, and Clarke MF (2003). Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci USA* **100**, 3983–3988.
[36] Ferrara N, Hillan KJ, Gerber HP, and Novotny W (2004). Discovery and development of bevacizumab, an anti-VEGF antibody for treating cancer. *Nat Rev Drug Discov* **3**, 391–400.
[37] Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey S, et al. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* **98**, 10869–10874.
[38] van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, et al. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536.
[39] Chang HY, Nuyten DS, Sneddon JB, Hastie T, Tibshirani R, Sorlie T, Dai H, He YD, van't Veer LJ, Bartelink H, et al. (2005). Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc Natl Acad Sci USA* **102**, 3738–3743.
[40] Liu R, Wang X, Chen GY, Dalerba P, Gurney A, Hoey T, Sherlock G, Lewicki J, Shedden K, and Clarke MF (2007). The prognostic role of a gene signature from tumorigenic breast-cancer cells. *N Engl J Med* **356**, 217–226.
[41] Reya T, Morrison SJ, Clarke MF, and Weissman IL (2001). Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111.
[42] Fillmore CM and Kuperwasser C (2008). Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Res* **10**, R25.
[43] Kok M, Koornstra RH, Margarido TC, Fles R, Armstrong NJ, Linn SC, Van't Veer LJ, and Weigelt B (2009). Mammosphere-derived gene set predicts outcome in patients with ER-positive breast cancer. *J Pathol* **8**, 316–326.
[44] Honeth G, Bendahl PO, Ringner M, Saal LH, Gruvberger-Saal SK, Lovgren K, Grabau D, Ferno M, Borg A, and Hegardt C (2008). The $CD44^+/CD24^-$ phenotype is enriched in basal-like breast tumors. *Breast Cancer Res* **10**, R53.
[45] Charafe-Jauffret E, Ginestier C, Iovino F, Wicinski J, Cervera N, Finetti P, Hur MH, Diebel ME, Monville F, Dutcher J, et al. (2009). Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. *Cancer Res* **69**, 1302–1313.
[46] Kaplan RN, Riba RD, Zacharoulis S, Bramley AH, Vincent L, Costa C, MacDonald DD, Jin DK, Shido K, Kerns SA, et al. (2005). VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. *Nature* **438**, 820–827.
[47] Calabrese C, Poppleton H, Kocak M, Hogg TL, Fuller C, Hamner B, Oh EY, Gaber MW, Finklestein D, Allen M, et al. (2007). A perivascular niche for brain tumor stem cells. *Cancer Cell* **11**, 69–82.
[48] Chi JT, Chang HY, Haraldsen G, Jahnsen FL, Troyanskaya OG, Chang DS, Wang Z, Rockson SG, van de Rijn M, Botstein D, et al. (2003). Endothelial cell diversity revealed by global expression profiling. *Proc Natl Acad Sci USA* **100**, 10623–10628.
[49] Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K, Chi JT, van de Rijn M, Botstein D, and Brown PO (2004). Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. *PLoS Biol* **2**, E7.
[50] Bogaerts J, Cardoso F, Buyse M, Braga S, Loi S, Harrison JA, Bines J, Mook S, Decker N, Ravdin P, et al. (2006). Gene signature evaluation as a prognostic tool: challenges in the design of the MINDACT trial. *Nat Clin Pract Oncol* **3**, 540–551.

IGF-I induced genes in stromal fibroblasts predict the clinical outcome of breast and lung cancer patients

Michal Rajski, Rosanna Zanetti-Dällenbach, Brigitte Vogel, Richard Herrmann, Christoph Rochlitz and Martin Buess

BMC Medicine. 2010 Jan 5;8(1):1. [PMID: 20051100]

BMC Medicine

RESEARCH ARTICLE Open Access

IGF-I induced genes in stromal fibroblasts predict the clinical outcome of breast and lung cancer patients

Michal Rajski[1], Rosanna Zanetti-Dällenbach[4], Brigitte Vogel[1], Richard Herrmann[1,3], Christoph Rochlitz[1,3], Martin Buess[1,2*]

Abstract

Background: Insulin-like growth factor-1 (IGF-I) signalling is important for cancer initiation and progression. Given the emerging evidence for the role of the stroma in these processes, we aimed to characterize the effects of IGF-I on cancer cells and stromal cells separately.

Methods: We used an *ex vivo* culture model and measured gene expression changes after IGF-I stimulation with cDNA microarrays. *In vitro* data were correlated with *in vivo* findings by comparing the results with published expression datasets on human cancer biopsies.

Results: Upon stimulation with IGF-I, breast cancer cells and stromal fibroblasts show some common and other distinct response patterns. Among the up-regulated genes in the stromal fibroblasts we observed a significant enrichment in proliferation associated genes. The expression of the IGF-I induced genes was coherent and it provided a basis for the segregation of the patients into two groups. Patients with tumours with highly expressed IGF-I induced genes had a significantly lower survival rate than patients whose tumours showed lower levels of IGF-I induced gene expression (P = 0.029 - Norway/Stanford and P = 7.96e-09 - NKI dataset). Furthermore, based on an IGF-I induced gene expression signature derived from primary lung fibroblasts, a separation of prognostically different lung cancers was possible (P = 0.007 - Bhattacharjee and P = 0.008 - Garber dataset).

Conclusion: Expression patterns of genes induced by IGF-I in primary breast and lung fibroblasts accurately predict outcomes in breast and lung cancer patients. Furthermore, these IGF-I induced gene signatures derived from stromal fibroblasts might be promising predictors for the response to IGF-I targeted therapies.
See the related commentary by Werner and Bruchim: http://www.biomedcentral.com/1741-7015/8/2

Background

There is a considerable amount of evidence that the insulin-like growth factor (IGF) family is important for cancer development and progression and IGF signalling is known to involve complex regulatory networks with numerous interacting ligands, receptors and binding proteins [1,2]. IGF-I, the first ligand of the family, may act as a tissue growth factor in an autocrine or paracrine manner or as a circulating hormone [3]. An elevated IGF-I level in the plasma is linked to an increased risk of developing ductal carcinoma *in situ* of the breast, invasive breast cancer, colorectal cancer, prostate cancer and lung cancer [4-9].

IGF-I signalling is crucial for tumour progression because it is involved in cell proliferation, differentiation, migration and survival [2,3,10-13]. On the molecular level, IGF-I is one of the factors that enables cells to pass the G1-S checkpoint in the cell cycle [14]. Normal mammary epithelial cells can be maintained and will proliferate with IGF-I in serum free cell culture media, underscoring the IGF-I's importance for the growth of breast epithelial cells [15,16]. In combination with mammogenic hormones, IGF induces ductal growth in mammary gland explant cultures [17]. Furthermore, IGF-I and IGF-II can suppress apoptosis of mammary

* Correspondence: Martin.Buess@claraspital.ch
[1]Department of Biomedicine, University of Basel, Hebelstrasse 20, CH-4031 Basel, Switzerland

epithelial cells induced by serum withdrawal [12]. *In vivo*, the involution of mammary glands is delayed in mice over-expressing human IGF-I due to reduced alveolar apoptosis [18]. During mammary gland development, IGF-I synergizes with estrogen in terminal end bud formation [19]. Finally, both IGF-I and IGF-II provide cancer cells with radioprotection and resistance to chemotherapeutic agents [20,21].

Further highlighting the importance of the IGF-I axis, the IGF-I receptor (IGF-IR) is crucial in cancer development and progression. The IGF-IR was found to be over-expressed and highly activated in malignant breast tumours compared with normal breast tissue [22,23]. Patients bearing an oestrogen receptor negative breast tumour have a worse prognosis when their tumour is positive for IGF-IR [24]. The functional importance of IGF-IR has been shown *in vitro* by inhibiting the receptor signalling which results in cancer cell apoptosis. *In vivo*, the inhibition of IGF-IR signalling prevents tumour formation in nude mice [1,25]. Moreover, IGF-IR-deficient fibroblasts cannot be transformed by viral or cellular oncogenes [26], supporting the importance of IGF-IR signalling in tumourigenesis.

That IGFs are involved in breast cancer migration and invasion has been demonstrated using dominant-negative IGF-IR constructs in MDA-435 breast cancer cells *in vitro* and *in vivo* [27]. Another experiment revealed that IGF-I stimulates cell motility, but not proliferation, in MDA-231BO cells in which the predominant adaptor protein for IGF-IR is the insulin receptor substrate 2 (IRS2) instead of the insulin receptor substrate 1 (IRS1). Further evidence supporting the involvement of IGF-IR and IRS2 axis in motility and metastasis comes from *in vivo* data. The mating of mice expressing the PyV-MT (polyomavirus middle T) oncogene, which induces breast cancer, with IRS2 *null* animals instead of wild-type animals results in their offspring showing a decrease in the formation of metastasis [28]. Thus, IGF-I is emerging as an important factor in tumourigenesis as a cell death inhibitor and a proliferation enhancer. Its involvement in tumour progression, metastasis and resistance to anti-neoplastic therapies makes it a promising drug target which is currently being examined in numerous clinical trials [29].

So far, the attention on IGF-I has focused on mitogenic and tumourigenic signalling in cancer cells [9,29,30]. With the increasing knowledge of the role of the tumour stroma in cancer initiation and progression, the role of IGF-I signalling in the stroma is of equal interest. In tumours, most of the IGF-I mRNA is localized in the stromal cells [31], especially fibroblasts [32], whereas most of the IGF-IR mRNA is in the tumour cells [33] which indicates that IGF-I produced in the stroma influences the tumour cells. However, there is evidence that IGF-I also influences the stroma. Stromal cells respond to IGF-I stimulation with increased proliferation, as do fibroblasts [34,35] and microvascular endothelial cells [36].

In addition to the response of the tumour cells to IGF-I, we specifically focused on the response of the stromal cells to this growth factor. Bendall *et al.* recently showed that the IGF-IR axis is involved in the establishment of the stem cell niche [37]. Blocking IGF-II/IGF-IR reduces the survival and clonogenicity of human embryonic stem cells (ES). Similarly, IGF-II alone is sufficient to maintain human ES cells in culture. In this system, IGF-II was expressed by autologously human-ES-cell-derived fibroblast-like cells.

In our study, we explore the role that IGF-I stimulation plays in cancer and stroma cells. We study the molecular changes that occur in primary normal and cancer-associated fibroblasts when they are stimulated with IGF-I. Furthermore, we hypothesized that gene expression changes in this system might be of prognostic significance in human cancer.

In this report, we show that primary normal and carcinoma-associated breast fibroblasts are sensitive to IGF-I. In addition, fibroblasts of different origin show a unified response to IGF-I. We also demonstrate that genes up-regulated in primary breast and lung fibroblasts may have prognostic significance in human breast cancer and lung adenocarcinomas.

Results

Effects of IGF-I on gene expression in breast cancer cells and stromal fibroblasts

In order to characterize the effects of IGF-I on tumour and stromal cells, we stimulated pre-starved MCF-7 cells and CCL-171 fibroblasts with 50 ng/ml IGF-I (a concentration within the physiological range) for 24 h. We then profiled gene expression changes using human exonic evidence-based oligonucleotide (HEEBO) microarrays. After stimulation, total RNA was extracted and amplified using a modified Eberwine procedure. The amplified RNA was labelled with the fluorescent dye Cy5 and pooled with Cy3 labelled reference RNA [38] and then the pooled RNA was hybridized onto HEEBO microarrays. After hybridization and washing, the arrays were scanned on a fluorescent microscope scanner and the raw data files were stored in the Stanford Microarray Database [39].

In order to establish the system, we characterized the response of both cell types to IGF-I separately. In both cell types, we observed a remarkable change in the gene expression profile following IGF-I stimulation (Figure 1). As the interaction between IGF-I and stromal cells in the tumour microenvironment has not yet been studied, we first characterized the IGF-I-induced genes in

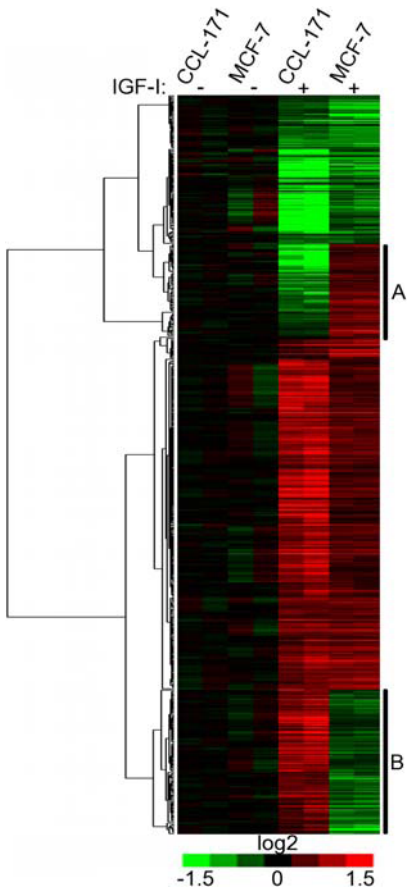


Figure 1 The effects of insulin-like growth factor-1 (IGF-I) on gene expression in CCL-171 fibroblasts and MCF-7 tumour cells. Unsupervised hierarchical clustering of genes deregulated in CCL-171 and MCF-7 cells upon IGF-I stimulation. The gene expression levels were normalized to the non-stimulated specimens as described. Genes are presented in rows and experiments in columns. The red and green colours provide information about up- or down-regulation, respectively. The intensity of the colour renders quantitative information about the change in expression level. IGF-I stimulation induces some common and some distinct effects on the gene expression profiles in different cell types. (A) Genes specifically up-regulated in MCF-7 cells involved in: epidermal growth factor and fibroblast growth factor signalling; protein metabolism and modification; nucleoside, nucleotide and nucleic acid metabolism. (B) Genes specifically up-regulated in CCL-171 cells include transcription factors and transferases, in addition to genes involved in Wnt and TGF-β signalling.

CCL-171 cells. The most prominent change after stimulation was a greater than 1.5-fold induction (mean: 2.35, standard deviation: 0.45) in the expression level of 370 genes (Additional file 1). The fibroblast derived IGF-I signature contains TTK, NEK2, PBK, SPBC24, RACGAP1, CLASP1, HECTD3, RCC2, MAD2L1, CDCA8, PTTG1, BIRC5, PKMYT1, HCAP-G, CCNB1, CENPF, CDC20, CKS2, SPAG5, PLK1, BUB1B, CCNF, KIF11, CDC25C, DLG7, BRRN1 and CDCA5, genes that are known to be involved in proliferation, cell cycle and mitotic cell division.

In order to check, in an unbiased way, what features the members of the IGF-I induced signature share and to verify the significance of enrichment of a specific gene ontology, we used the GO TermFinder tool [40]. The analysis revealed that the fibroblast derived IGF-I signature is significantly enriched for genes involved in biological processes such as M phase, mitotic cell cycle, mitosis and cell cycle, with a P value equal to ≈ 1.03e-10 and cell division with P ≈ 1.03e-8 (Table 1 and Additional file 2). In addition, among the 370 unique genes up-regulated by IGF-I we found genes that are involved in angiogenesis, the p53 pathway and integrin and Wnt signalling. The mRNA expression level of six soluble factors already recognized in cancer biology (POSTN, TNC, CSPG2, LOXL1, ATRN, FBS1) increases in response to IGF-I stimulation, suggesting that these factors play some role in stimulating tumour cell proliferation and metastasis.

We selected the MCF-7 cell line, the well-known representative of the luminal type breast cancer, to assess the global gene expression effects of IGF-I stimulation. IGF-I stimulation increased the mRNA expression level of numerous genes (such as BMP7, ID1, ID3, SRF and VEGF) that play a role in tumour biology as well as genes with ontologies assigned to protein metabolism (RPS6KA4, PSMC4, MAPK6, LMAN2L, RPL8, EIF5 and CEBPB), responses to a protein stimulus and to unfolded protein (HSP90AA1, HSPE1, HSPA1A, DNAJA1, HSPA4, HSP90AB1, HSPH1, and DNAJB1). In contrast to the upregulation of genes involved in the proliferation that we found in CCL-171 cells, the gene expression pattern in MCF-7 cells stimulated with IGF-I is not significantly associated with the cell cycle or proliferation.

After comparing these gene expression patterns, we hypothesized that the mesenchymal stromal cells and malignant epithelial cells exhibit distinct gene expression changes in response to IGF-I stimulation. In order to test this hypothesis, we compared the gene expression profiles of CCL-171 and MCF-7 cells with, and without, IGF-I stimulation. Within each cell type, we subtracted the expression profile of unstimulated cells from IGF-I stimulated cells and then filtered and merged the profiles and performed a hierarchical clustering of the genes. The results were visualized with a heat map using TreeView software [41], which showed that the IGF-I stimulus induced some common and some distinct effects on gene expression in the two cell types (Figure 1). This is easily explained when we consider the differences in the distinct default gene expression profiles of the two cell types, including the well-known markers for epithelial and mesenchymal cells

Table 1 Gene ontology terms for genes up-regulated in CCL-171 cells by insulin-like growth factor (IGF-I).

Gene ontology term	Cluster frequency	Gene frequency in background	Corrected *P*-value	FDR	False positives
M phase	42 out of 325 genes, 12.9%	67 out of 2133 genes, 3.1%	1.84E-16	0.0%	0.0
Cell cycle phase	45 out of 325 genes, 13.8%	82 out of 2133 genes, 3.8%	1.46E-14	0.0%	0.0
Cell division	38 out of 325 genes, 11.7%	69 out of 2133 genes, 3.2%	5.07E-12	0.0%	0.0
M phase of mitotic cell cycle	32 out of 325 genes, 9.8%	52 out of 2133 genes, 2.4%	1.06E-11	0.0%	0.0
Cell cycle process	50 out of 325 genes, 15.4%	113 out of 2133 genes, 5.3%	2.13E-11	0.0%	0.0
Nuclear division	31 out of 325 genes, 9.5%	51 out of 2133 genes, 2.4%	4.65E-11	0.0%	0.0
Mitosis	31 out of 325 genes, 9.5%	51 out of 2133 genes, 2.4%	4.65E-11	0.0%	0.0
Cell cycle	57 out of 325 genes, 17.5%	143 out of 2133 genes, 6.7%	7.11E-11	0.0%	0.0
Organelle fission	31 out of 325 genes, 9.5%	52 out of 2133 genes, 2.4%	9.95E-11	0.0%	0.0
Mitotic cell cycle	39 out of 325 genes, 12.0%	86 out of 2133 genes, 4.0%	7.94E-09	0.0%	0.0
Organelle organization	67 out of 325 genes, 20.6%	217 out of 2133 genes, 10.2%	4.21E-07	0.0%	0.0
Cellular component organization	91 out of 325 genes, 28.0%	358 out of 2133 genes, 16.8%	1.78E-05	0.0%	0.0
Microtubule-based process	21 out of 325 genes, 6.5%	40 out of 2133 genes, 1.9%	2.56E-05	0.0%	0.0
Regulation of mitotic cell cycle	17 out of 325 genes, 5.2%	31 out of 2133 genes, 1.5%	2.40E-04	0.0%	0.0
Spindle organization	10 out of 325 genes, 3.1%	14 out of 2133 genes, 0.7%	2.70E-03	0.4%	0.1
Microtubule-based Movement	10 out of 325 genes, 3.1%	15 out of 2133 genes, 0.7%	0.01	0.6%	0.1
Mitotic cell cycle checkpoint	9 out of 325 genes, 2.8%	13 out of 2133 genes, 0.6%	0.01	0.6%	0.1
Regulation of cell cycle	27 out of 325 genes, 8.3%	79 out of 2133 genes, 3.7%	0.01	0.6%	0.1
Biological regulation	152 out of 325 genes, 46.8%	785 out of 2133 genes, 36.8%	0.04	0.6%	0.1
Microtubule cytoskeleton organization	12 out of 325 genes, 3.7%	24 out of 2133 genes, 1.1%	0.05	0.7%	0.1

Detailed list of gene ontology terms specifically up-regulated in CCL-171 fibroblasts upon IGF-I stimulation in comparison to background file including all genes deregulated by IGF-I. Bonferroni corrected *p* values for of the output from GO::Termfinder for process ontology are listed.
FDR, false discovery rate.

(Additional file 3). The two gene clusters with discordant gene expression changes (Figure 1) were examined with the GO TermFinder tool. Genes that are up-regulated in CCL-171 and down-regulated in MCF-7 cells belong to the following ontologies: Wnt and TGF-β signalling and nucleic acid binding (transcription factors and transferases). Genes that are up-regulated in MCF-7 and down-regulated in CCL-171 cells are involved in protein metabolism and modification, as well as nucleoside, nucleotide and nucleic acid metabolism. This list also contains genes involved in epidermal growth factor and fibroblast growth factor signalling.

Thus, we concluded that, when epithelial and mesenchymal cells are exposed to IGF-I, they show some concordant and some discordant gene expression changes. Based on these observations, and given that the role of IGF-I in the stromal compartment is not yet well characterized, we further focused on the IGF-I response in stromal cells.

In order to characterize the gene expression programme induced in fibroblasts upon IGF-I stimulation, we extended our survey to primary breast fibroblasts. Rinn *et al.* [42] reported that fibroblasts from different body sites have unique default gene expression profiles, which led us to believe that distinct fibroblasts may respond differently to stimulation with IGF-I. Therefore, we felt that it would be important to analyse primary breast fibroblasts from breast cancer patients in order to examine the role that these stromal cells play in breast cancer.

After obtaining informed consent from three patients with oestrogen and progesterone receptor positive and HER-2/neu negative invasive ductal adenocarcinoma of the breast, tissue specimens were obtained during breast tumour excision. An experienced breast pathologist distinguished tumour tissue from adjacent normal tissue. Carcinoma associated fibroblasts (CAF) and normal fibroblasts obtained from normal breast tissue of the same patient were cultured separately. The desired purity of the cell culture was obtained by serial passaging and separation with magnetic beads targeting fibroblast-specific antigens. Both cell types, CAF and normal fibroblasts, were stimulated with IGF-I and gene expression profiles were observed. We confirmed that the profiled cells were, indeed, mesenchymal fibroblasts because they showed an elevated expression of fibroblast markers, such as fibronectin 1 (FN1) and cadherin 2 (CDH2), and lacked E-cadherin (CDH1) expression (Additional file 4). The expression level of these specific markers did not change upon IGF-I stimulation (data not shown). All of the primary fibroblasts had a slightly higher IGF-IR mRNA expression level (mean: 1.6-fold; standard

deviation: 0.24) compared to reference mRNA isolated from a pool of 11 cell lines [38], indicating that they might be responsive to IGF-I stimulation. The IGF-IR mRNA expression level decreased after IGF-I stimulation (mean: 0.6-fold; standard deviation: 0.09). In order to systematically identify significant gene expression changes upon IGF-I stimulation in primary cells, we applied a two-class significance analysis of microarray (SAM) data [43]. One class was formed by fibroblasts starved in low serum medium and the other class consisted of the same cells stimulated with IGF-I. SAM analysis revealed 208 gene IDs up-regulated and 300 gene IDs significantly repressed in stimulated cells (false discovery rate [FDR] ≤ 0.05%, Additional file 5). The 208 up-regulated genes were used to create the breast fibroblast derived IGF-I signature (Figure 2A). By comparing the gene function with the GO Termfinder, we observed that the genes up-regulated by IGF-I in primary breast fibroblasts (Additional file 5) share similar features to those up-regulated in IGF-I-stimulated CCL-171 cells (Additional file 1), suggesting that they are involved in the same processes (proliferation, cell cycle and mitosis - Additional file 6, Table 2). Contrary to our expectations, we did not find any significant differences in the response to IGF-I between CAF and normal fibroblasts.
Taken together, primary fibroblasts coming from breast cancer and the normal breast, as well as CCL-171 fibroblasts, respond to IGF-I stimulation and display up-regulation of similar gene signatures involved in cell proliferation and mitotic cell division.

In order to verify that the gene expression profile is reflected by a phenotypic alteration upon IGF-I stimulation, we examined the proliferation rate of the fibroblasts. The cells were seeded and starved in low serum medium for 48 h in order to exclude any effects of fetal bovine serum (FBS) from regular cell growth culture conditions. The cells were then stimulated with IGF-I, and the cell proliferation was assessed with a colourimetric method using WST-1. Primary breast fibroblasts (Figure 2C), both normal and CAF, grew significantly faster when stimulated with IGF-I rather than unstimulated cells ($P < 0.0001$ for 24, 48 and 72 h, *t*-test, two-sided; $P < 0.0001$, analysis of variance [ANOVA]). A similar response to IGF-I stimulation was observed in CCL-171 fibroblasts as presented in Figure 2B ($P < 0.0001$ for 24, 48 and 72 h, *t*-test, two-sided; $P < 0.001$, ANOVA).

Taken together, these suggest that primary fibroblasts stimulated with IGF-I induce expression of genes involved in cell proliferation and mitotic cell division.

Relevance of the fibroblast derived IGF-I induced gene signature *in vivo*

In order to verify the relevance of our *in vitro* experiments, we checked the expression of the breast fibroblast derived IGF-I signature in microarray data of early stage breast cancer biopsies from 295 patients from the Netherlands Cancer Institute (NKI), which are publicly available [44]. In the NKI dataset, the expression of the genes belonging to breast fibroblast derived IGF-I signature was coherent, providing a basis for segregation of the tumours into two groups. In one group the signature was up-regulated and in the other group the signature was down-regulated (left and right side of Figure 3A, respectively). As visualized with Kaplan-Meier plots (Figure 3B), the patients with early stage breast cancers with a high expression level of the breast fibroblast derived IGF-I signature had a significantly higher risk of developing metastasis than the patients with a low expression level (P = 6.75e-05, 52% versus 73% after 10 years, hazard ratio (HR): 2.24, 95% confidence interval [CI]: 1.5-3.4; top panel). In parallel, the overall survival rate was significantly lower for patients with up-regulation of the breast fibroblast derived IGF-I signature (P = 7.96e-09, 55% versus 86% after 10 years, HR: 4.03, CI: 2.4-6.8; middle panel). The same coordinated behaviour and segregation of tumours could also be observed in a set of advanced breast cancers from Norway/Stanford [45,46]. In a univariate analysis, patients with high expression levels of IGF-I induced genes had a significantly shorter disease-specific survival than patients with low expression levels (P = 0.0219, HR: 2.6, CI: 1.1-6.2, data not shown). In addition, as the classification of data based on hierarchical clustering was suggested to be unstable and codependent on many factors like presence of missing values [47], we validated the results using continuous scoring and stratified the patients of the NKI dataset based on a score derived from the average expression level of breast fibroblast derived IGF-I signature. In agreement with the results obtained by hierarchical clustering, the continuous scoring divided the early breast cancer patients (NKI dataset) [44] into two groups with significantly different outcomes (distant metastasis-free survival: P = 3.6e-07 and overall survival: P = 3.5e-09; Additional file 7).
On the molecular level, an interaction of IGF-I with the oestrogen receptor (ER) has been described. Therefore, we performed a multivariable analysis corrected for ER status (positive versus negative) in the early and advanced breast cancer datasets. The breast fibroblast derived IGF-I signature was able to stratify breast cancer patients into groups with significantly different outcomes even when corrected for ER status. The results of the multivariable analysis were significant (overall survival: P = 1.6e-09, time to metastasis: P = 2.2e-4 in the NKI dataset and disease specific survival in the Norway/Stanford dataset P = 8.6e-5, respectively). In both datasets, the combination of ER negative receptor status and up-regulation of the breast fibroblast derived IGF-I

Table 2 Gene ontology terms for genes up-regulated in breast fibroblasts by insulin-like growth factor-1 (IGF-I).

Gene Ontology term	Cluster frequency	Gene frequency in background	Corrected *P*-value	FDR	False positives
M phase	36 out of 186 genes, 19.4%	175 out of 8918 genes, 2.0%	2.00E-23	0%	0.00
Cell cycle phase	37 out of 186 genes, 19.9%	223 out of 8918 genes, 2.5%	1.18E-20	0%	0.00
Cell cycle process	41 out of 186 genes, 22.0%	314 out of 8918 genes, 3.5%	4.69E-19	0%	0.00
Nuclear division	28 out of 186 genes, 15.1%	128 out of 8918 genes, 1.4%	1.64E-18	0%	0.00
Mitosis	28 out of 186 genes, 15.1%	128 out of 8918 genes, 1.4%	1.64E-18	0%	0.00
Cell cycle	46 out of 186 genes, 24.7%	425 out of 8918 genes, 4.8%	2.85E-18	0%	0.00
M phase of mitotic cell cycle	28 out of 186 genes, 15.1%	131 out of 8918 genes, 1.5%	3.21E-18	0%	0.00
Organelle fission	28 out of 186 genes, 15.1%	133 out of 8918 genes, 1.5%	4.99E-18	0%	0.00
Mitotic cell cycle	33 out of 186 genes, 17.7%	229 out of 8918 genes, 2.6%	3.30E-16	0%	0.00
Cell division	28 out of 186 genes, 15.1%	164 out of 8918 genes, 1.8%	1.88E-15	0%	0.00
Microtubule-based process	24 out of 186 genes, 12.9%	130 out of 8918 genes, 1.5%	8.42E-14	0%	0.00
Microtubule-based movement	13 out of 186 genes, 7.0%	47 out of 8918 genes, 0.5%	3.61E-09	0%	0.00
Spindle organization	10 out of 186 genes, 5.4%	27 out of 8918 genes, 0.3%	3.73E-08	0%	0.00
Cytoskeleton-dependent intracellular transport	13 out of 186 genes, 7.0%	56 out of 8918 genes, 0.6%	4.11E-08	0%	0.00
Organelle organization	44 out of 186 genes, 23.7%	737 out of 8918 genes, 8.3%	6.24E-08	0%	0.00
Chromosome segregation	10 out of 186 genes, 5.4%	41 out of 8918 genes, 0.5%	3.83E-06	0%	0.00
Microtubule cytoskeleton organization	12 out of 186 genes, 6.5%	70 out of 8918 genes, 0.8%	8.84E-06	0%	0.00
Phosphoinositide-mediated signalling	8 out of 186 genes, 4.3%	27 out of 8918 genes, 0.3%	2.32E-05	0%	0.00
Mmitotic sister chromatid segregation	7 out of 186 genes, 3.8%	22 out of 8918 genes, 0.2%	9.41E-05	0%	0.00
Sister chromatid segregation	7 out of 186 genes, 3.8%	22 out of 8918 genes, 0.2%	9.41E-05	0%	0.00
Cellular component organization	50 out of 186 genes, 26.9%	1187 out of 8918 genes, 13.3%	3.40E-04	0%	0.00
Regulation of mitotic cell cycle	10 out of 186 genes, 5.4%	72 out of 8918 genes, 0.8%	1.03E-03	0%	0.00
Second-messenger-mediated signalling	8 out of 186 genes, 4.3%	44 out of 8918 genes, 0.5%	1.36E-03	0%	0.00
Regulation of cell cycle	16 out of 186 genes, 8.6%	197 out of 8918 genes, 2.2%	1.68E-03	0%	0.00
Cytoskeleton organization	16 out of 186 genes, 8.6%	198 out of 8918 genes, 2.2%	1.80E-03	0%	0.00
Protein polymerization	6 out of 186 genes, 3.2%	23 out of 8918 genes, 0.3%	2.63E-03	0%	0.00
Positive regulation of mitosis	4 out of 186 genes, 2.2%	7 out of 8918 genes, 0.1%	2.73E-03	0%	0.00
Amino acid biosynthetic process	6 out of 186 genes, 3.2%	25 out of 8918 genes, 0.3%	4.45E-03	0%	0.00
Chromosome localization	4 out of 186 genes, 2.2%	8 out of 8918 genes, 0.1%	0.01	0%	0.00
Establishment of chromosome localization	4 out of 186 genes, 2.2%	8 out of 8918 genes, 0.1%	0.01	0%	0.00
Cell cycle checkpoint	7 out of 186 genes, 3.8%	47 out of 8918 genes, 0.5%	0.02	0%	0.04
Establishment of localization in cell	24 out of 186 genes, 12.9%	472 out of 8918 genes, 5.3%	0.02	0%	0.04
Cellular localization	25 out of 186 genes, 13.4%	507 out of 8918 genes, 5.7%	0.03	0%	0.04
DNA metabolic process	17 out of 186 genes, 9.1%	275 out of 8918 genes, 3.1%	0.03	0%	0.06
Amine biosynthetic process	6 out of 186 genes, 3.2%	35 out of 8918 genes, 0.4%	0.03	0%	0.08
Serine family amino acid biosynthetic process	3 out of 186 genes, 1.6%	5 out of 8918 genes, 0.1%	0.04	0%	0.08

Detailed information and *P* values for single biological processes up-regulated by IGF-I in primary breast fibroblasts in comparison to background file including all genes used for significant analysis of mcroarray analysis as revealed by gene ontology term finder tool with a Bonferoni corrected *P* value higher than 0.05. FDR, false discovery rate.

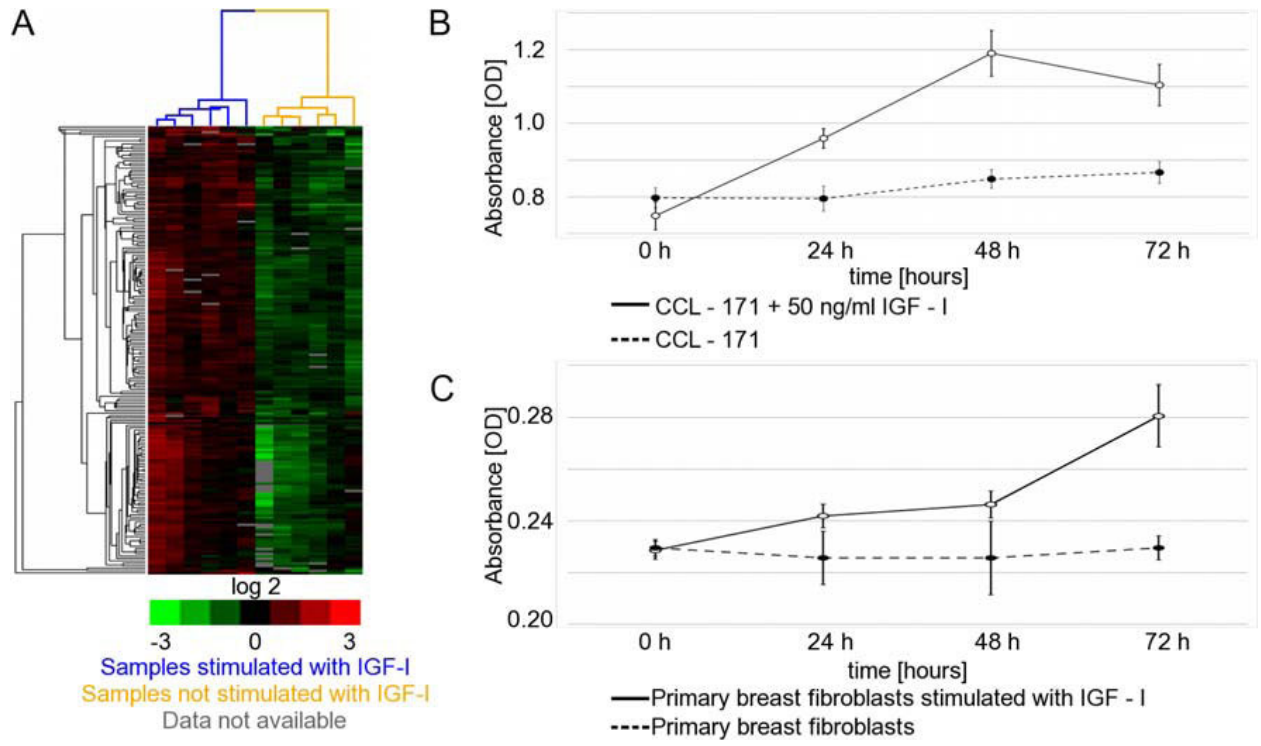


Figure 2 Effects of insulin-like growth-1 (IGF-I) stimulation on primary breast fibroblasts and CCL-171 fibroblasts. (A) Unsupervised hierarchical clustering of genes differentially expressed in fibroblasts upon IGF-stimulation. Unsupervised hierarchical clustering of genes differentially expressed between IGF-I stimulated and non-stimulated primary breast fibroblasts as discovered by SAM (genes with a false discovery rate ≤ 0.05% are represented). Grey fields indicate missing expression values. The colour of dendrogram branches renders information about sample stimulation; yellow = not stimulated and blue = stimulated with IGF-I (50 ng/mL). (B) IGF-I induced proliferation of CCL-171 cells. Cell proliferation assay based on absorbance measurement of WST-1. Formazan absorbance correlates to the cell number. Average absolute absorbance of replicates of CCL-171 cells stimulated with 50 ng/mL IGF-I in comparison to non-stimulated cells at different time points. Points represent the average of six replicates per condition and correspond to the cell number. The vertical error bars denote the standard deviation. Stimulation of CCL-171 cells with IGF-I induces significant, constant cell growth after 24, 48 and 72 h. (C) IGF-I induced proliferation of primary breast fibroblasts. Cell proliferation assay based on absorbance measurement of WST-1. Points represent the average absolute absorbance of a minimum of eight replicates of six primary fibroblasts (carcinoma associated fibroblasts and normal fibroblasts) after 24, 48 and 72 h. Error bars correspond to the magnitude of the standard deviation. Stimulation of primary breast fibroblasts with IGF-I induces significant, constant cell growth.

signature had the worst outcome. Additionally, in early stage breast cancer, the breast fibroblast derived IGF-I signature was able to segregate ER positive breast cancer patients into two groups with significantly different outcomes (P = 1.6e-5, Figure 3B, lowest panel). In summary, we found that genes induced in primary breast fibroblasts upon IGF-I stimulation predict the outcome of breast cancer patients. Furthermore, the expression signature distinguishes between patients with ER positive cancer who have significantly different prognoses.

Correlation of the IGF-I induced gene signature with previously published prognostic gene expression signatures

As the breast fibroblast derived IGF-I signature is a prognostic marker in human breast cancer, we next sought to see if the signature might be related to other previously published gene-expression signatures, which were useful prognosticators in the NKI dataset. To this aim, we correlated the signatures based on their centroids, which represent the average expression values of all genes building the signature in a single tumour specimen, using the Pearson correlation test. First, we checked the correlation of the breast fibroblast derived IGF-I signature centroid to the wound signature centroid [48], which was created based on the response of fibroblasts to serum stimulation. The breast fibroblast derived IGF-I signature, as presented in Figure 4, was highly correlated to the wound signature (0.76). It was also moderately correlated (0.69) to basal type breast cancer [46]. Furthermore, the breast fibroblast derived IGF-I signature was highly reverse-correlated to the good-risk 70-genes signature (-0.74) [49]. The good-risk70-genes signature was created n order to predict freedom from metastasis in this same dataset. The

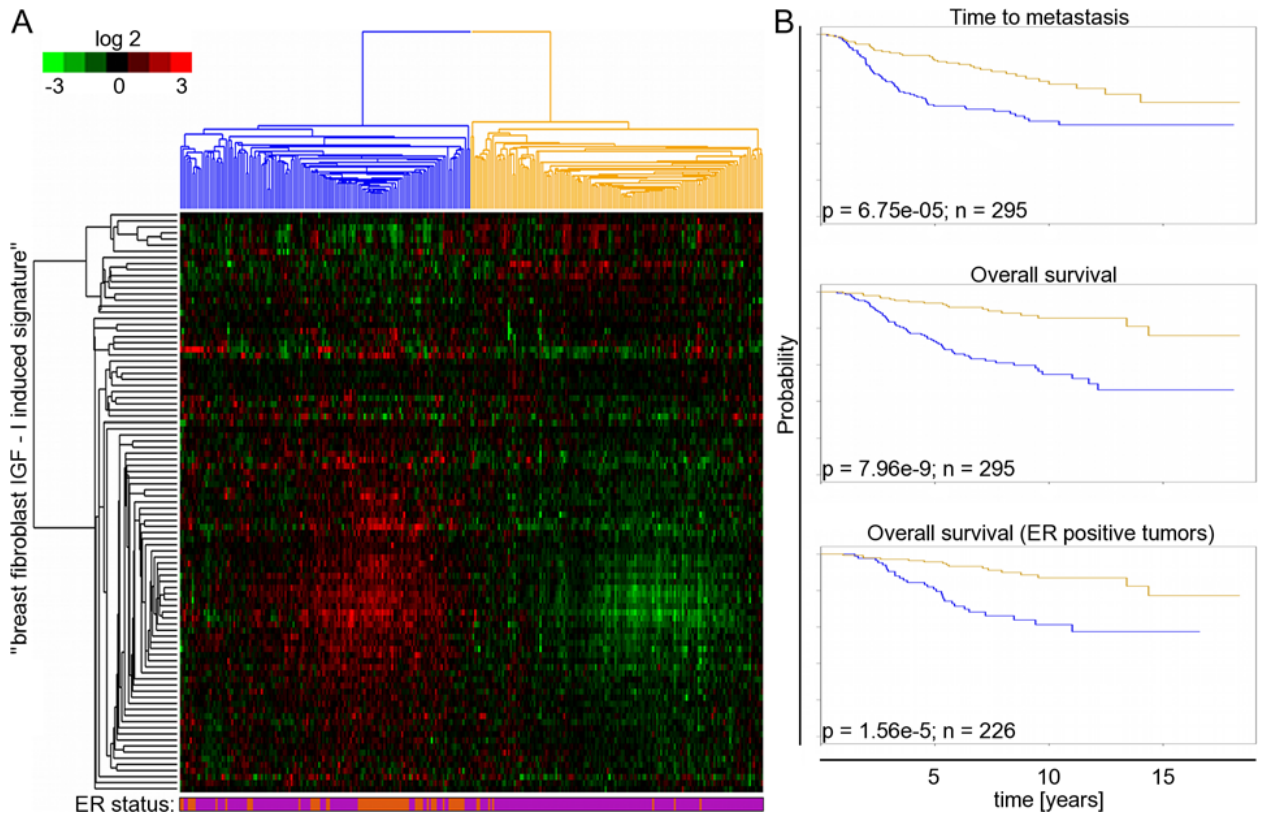


Figure 3 Breast fibroblast derived IGF-I signature in early stage breast cancer. (A) Unsupervised hierarchical clustering of breast fibroblast derived IGF-I signature in Netherlands Cancer institute dataset. The expression values of genes in the breast fibroblast derived IGF-I signature revealed by signficant analysis of microarray were extracted from a published expression study of 295 early stage breast cancers from the Netherlands Cancer Institute (NKI). Genes are presented in rows and experiments in columns. Breast fibroblast derived IGF-I signature stratifies early breast cancer patients (NKI) into two groups with high (blue) or low (yellow) expression levels of genes representing the signature. Horizontal bar below the figure represents positive (purple) or negative (orange) ER status. (B) Relationship of expression level of genes building breast fibroblast derived IGF-I signature with distant metastasis free and overall survival. Kaplan-Meier curves representing the clinical outcomes of tumors exhibiting high (blue curve) and low (yellow curve) expression levels of the IGF-I induced signature. The upper two figures represent all patients and the bottom figure shows only patients with oestrogen receptor positive breast tumours.

detailed list of correlation values for all of the signatures may be found in Additional file 8.

IGF induced genes are prognostic in lung cancer

Knowing that the gene expression signature derived from primary breast fibroblasts in response to IGF-I stimulation is relevant *in vivo*, and is a strong prognostic factor in human breast cancer, we investigated this finding to see if it could be generalized to other types of human cancer. We felt that this was likely because of the similarity between the IGF-I responses of primary breast fibroblasts and CCL-171 lung fibroblasts. We decided to check our hypothesis using the IGF-I derived signature from CCL-171 *in vitro* in lung cancer datasets. Global gene expression profiles of 67 human lung cancers were derived from 56 patients; 24 had survival data published by Garber *et al.* [50] (GEO: GSE3398). As shown in Figure 5A, in this dataset the expression of the lung fibroblast derived IGF-I gene signature was clear, even though the expression data for many genes was missing. This provided a basis for segregation of the tumours into two groups. The two groups were described as having the core part (proliferation associated genes) of the signature up-regulated or down-regulated (left and right side of Figure 5A, respectively). As visualized by Kaplan-Meier plots (Figure 5B), the patients with high expression levels of IGF-I induced genes had a significantly shorter overall survival (P = 0.008; n = 24, 60% versus 0% after 2 years, HR: 7.74, CI: 1.9-31.6). Thus, we concluded that the lung fibroblast derived IGF-I signature is a prognostic marker in lung cancer.

We then decided to validate our findings in a larger and better-annotated dataset published by Bhattacharjee [51], which contains microarray profiles of 203 tumours with clinical annotation for 125 of them. In line with our hypothesis, the expression of the lung fibroblast derived IGF-I signature was coherent, providing a basis

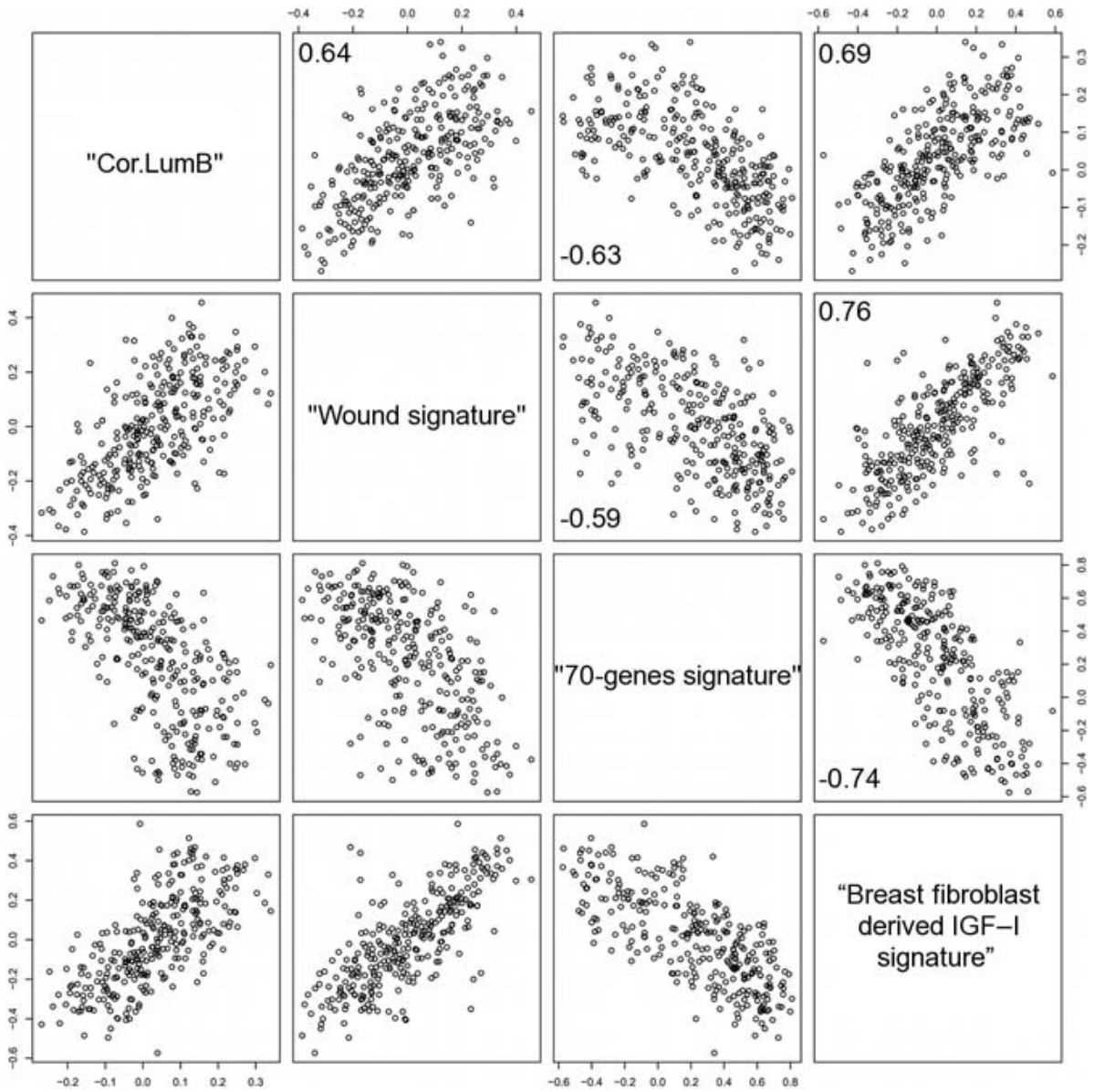


Figure 4 Correlation of the breast fibroblast derived insulin-like growth factor-I (IGF-I) signature with previously reported prognosticators in breast cancer. Correlation of the good-risk 70-genes signature centroid [49], the wound signature centroid [60], the basal type of breast cancer created by Soerlie [46] and the breast fibroblast IGF-I induced signature score in the Netherlands Cancer Institute dataset. Pairwise scatterplot-matrix of four gene signatures. Pearson correlations for the signature are shown in the corners of the plots.

for segregation of the tumours into two groups. The patients with a high expression level of the signature (right side of Figure 5C) had a significantly shorter disease-free survival (P = 0.001; 45% versus 82% after 5 years, HR: 3.7, CI: 1.5-9.4) and overall survival (P = 0.007; 38% versus 66% after 5 years, HR: 2.1, CI: 1.2-3.8) than the patients with low expression of the signature (left side of Figure 5C). Both Kaplan-Meier curves are shown in Figure 5D.

Taken together, these findings indicate that genes induced by IGF-I in human lung fibroblasts are helpful in predicting outcomes in human lung cancer. As the signature derived from breast and lung fibroblasts upon IGF-I stimulation is a prognostic marker for lung cancer, we suggest that the response of stromal fibroblasts to IGF-I might be a universal feature of cancer.

In order to further validate the general effect of IGF-I on fibroblasts, the ability of the breast and the lung

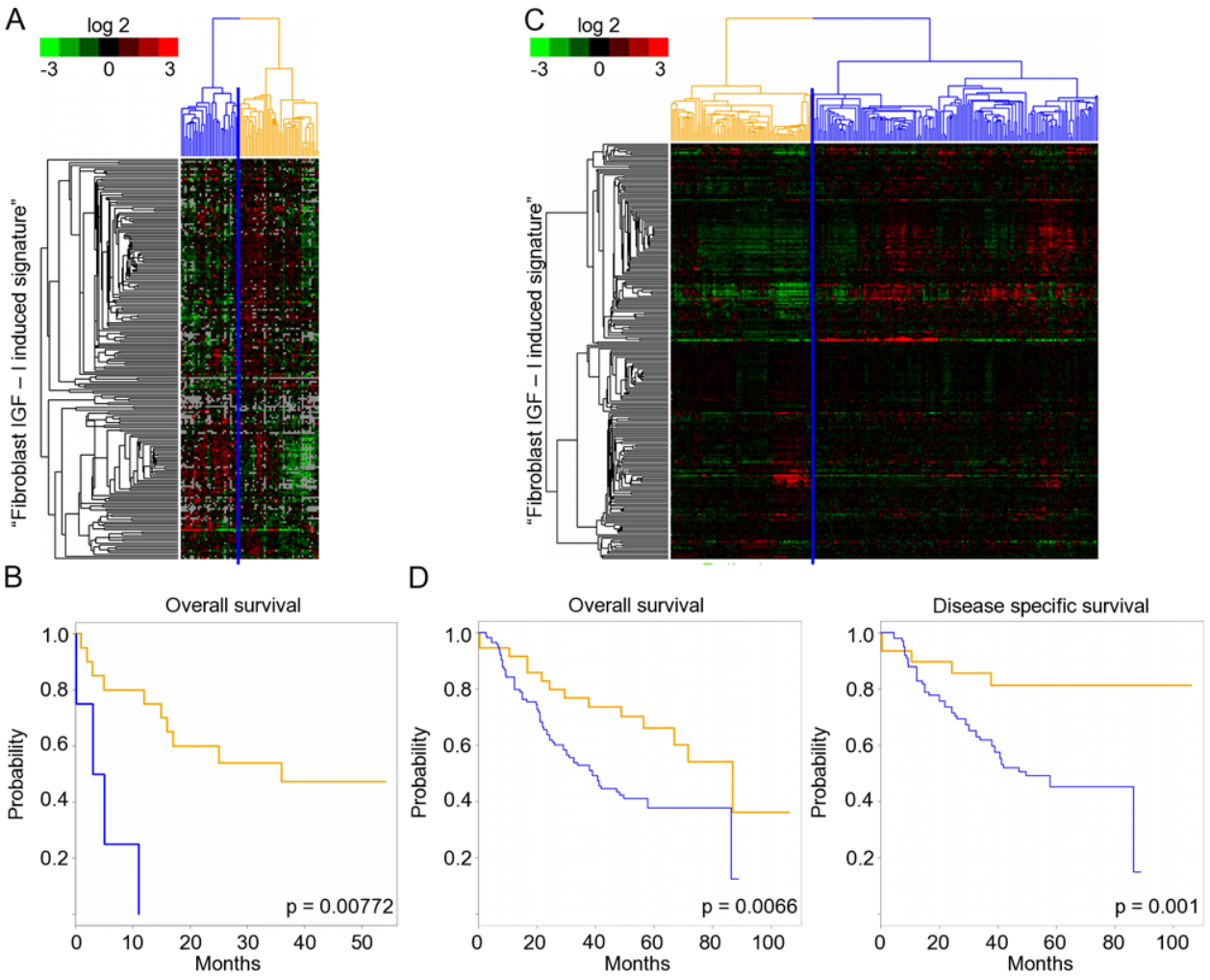


Figure 5 Fibroblast derived insulin-like growth factor-I (IGF-I) signature divides lung cancer patients into two groups with significantly different outcome. (A) Unsupervised hierarchical clustering of fibroblast derived IGF-I signature in Garber lung cancer dataset. The expression values of genes in the fibroblast derived IGF-I signature were extracted from a published expression study by Garber [50]. Genes are presented as rows and the experiments are presented as columns. Although some gene expression data are missing, the fibroblast derived IGF-I signature stratifies lung cancer patients into two groups with high (blue) or low (yellow) expression levels of genes representing the signature. (B) Relationship of expression level of genes building fibroblast derived IGF-I signature with overall survival in Garber data. Kaplan-Meier curves denoting the clinical outcomes of the indicated tumours exhibiting high (blue curve) and low (yellow curve) expression levels of the signature. (C) Unsupervised hierarchical clustering of fibroblast derived IGF-I signature in Bhattacharjee lung cancer dataset. The expression values of genes in the fibroblast derived IGF-I signature were extracted from a published expression study by Bhattacharjee [51]. Genes are presented as rows and the experiments are presented as columns. Fibroblast derived IGF-I signature stratifies adenocarcinoma patients into two groups with high (blue) or low (yellow) expression levels of genes representing the signature. (D) Relationship of expression level of genes building fibroblast derived IGF-I signature with overall survival and disease specific survival in Bhattacharjee dataset. Kaplan-Meier curves illustrating the clinical outcomes of the indicated tumours exhibiting high (blue curve) and low (yellow curve) expression levels of the signature.

fibroblast derived IGF-I signatures to be a prognostic factor in a non-site matching dataset was crosschecked. The breast fibroblast derived IGF-I signature was able to stratify patients with lung cancer (Bhattacharjee dataset) into two groups with significantly different rates of survival (overall survival with P = 0.043 and disease free survival with P = 0.022, data not shown). Also in this dataset, we validated the results obtained by hierarchical clustering using a continuous score and found a significant correlation (overall survival P = 0.005 and disease specific survival P = 0.001; Additional file 9).

The lung fibroblast derived IGF-I signature was also able to arrange breast carcinoma patients from the NKI dataset into two groups with significantly different times for metastasis free survival and overall survival (P = 7.9e-9 and P = 9.8e-6, respectively, data not shown).

N order to further cross validate the IGF-I signatures derived from fibroblasts of different origins, the

correlation of the centroids for the signature obtained from lung fibroblasts were correlated to the signature derived from human primary breast fibroblasts (0.77; *P* - value < 2.2e-16, Additional file 10) in the NKI breast cancer dataset. The strong and significant correlation supports their similarity.

Discussion

IGF-I has multiple effects on tumour initiation, development and progression and its effects on the cancer cells have been well described [13]. However, solid tumours do not consist only of malignant epithelial cells; rather, they form organ-like structures with a stroma consisting of fibroblasts, inflammatory cells and endothelial cells. Therefore, an endocrine or paracrine stimulus such as IGF-I might influence both the tumour cells and the stromal cells. The goal of this study was to characterize the effects of IGF-I on the cancer cells and the stromal fibroblasts in parallel. On the molecular level, cancer cells and fibroblasts show distinct response patterns to stimulation with IGF-I (Figure 1), including differential expression of genes involved in proliferation, protein metabolism and Wnt and TGF-β signalling. Focusing on the effect of IGF-I on MCF-7 cells, we observed alterations in protein metabolism. Similar changes in protein metabolism, including up-regulation of genes involved in transport and biosynthesis of amino acids, had already been reported previously in a global gene expression study of MCF-7 cells endogenously overexpressing IGF-I [52]. Additionally, we noted an up-regulation of VEGF in MCF-7 cells treated with IGF-I. VEGF is a known target gene for IGF signalling [52], with well-described implications in tumour progression and dissemination. Similar to our results, up-regulation of genes involved in metabolism and biosynthesis have been described in a comparable system of MCF-7 cells stimulated with exogenous IGF-I [53]. Apart from the similarities to the study by Creighton *et al.* [53], we also found discrepancies in the gene expression profile of proliferation-associated genes. The main reason might be because of the gene-wise standardization of the unstimulated samples applied in our setup which eliminated the inherent pattern of proliferation in MCF-7 cells.

With a focus on the stroma, there are studies showing that human dermal fibroblasts [35] and IMR90 fibroblasts [34] respond to IGF-I stimulation. Furthermore, it has been shown that primary breast fibroblasts overexpress IGF-I and IGF-II (normal and malignant derived fibroblasts, respectively) [31,54] but none of these studies focused on the effects of IGF-I signalling on global gene expression. There was only one small study with first generation microarrays profiling the global gene expression effects of IGF-I stimulation in NIH-3T3 mouse fibroblasts, which showed an up-regulation of proliferation-associated genes [55]. To the best of our knowledge, we are the first to show microarray gene expression profiles of primary human breast fibroblasts in response to IGF-I. The gene expression changes induced by IGF-I in fibroblasts contained several soluble factors, such as POSTN, which was reported to be involved in bone metastasis formation and angiogenesis [56,57], TNC, which enhances tumour cell proliferation [58], as well as LOXL1, a member of lysyl oxidase family, similar to LOXL2, that might act on or in the vicinity of epithelial cells during tissue remodelling. LOXL2 has previously been reported to be involved in an invasiveness process [59] and specifically expressed by fibroblasts in tumour tissue [60]. The presence of these factors indicates that the IGF-I activated stroma enhances proliferation and the metastatic potential of the cancer cells.

That one single stimulus has both common and distinct effects on cells of different origins has been shown previously on the global gene expression scale for the response to oxygen deprivation under hypoxic conditions [61]. To the best of our knowledge, our experiments are the first to make a direct comparison of the effects of IGF-I on different cell types. Most interestingly, among the genes that were upregulated only in CCL-171 cells, and not in MCF-7 cells on IGF-I stimulation, did we observe many transcription factors (FUBP3, TEAD2, KLF16, SP3 and PIK3R3 involved in insulin receptor signalling pathway, MKNK2, SH3BP2 and CIT) all taking part in cell surface receptor linked signal transduction. Stromal cell specific genes among the IGF-I induced genes were of interest when we correlated this signature with *in vivo* data derived from whole tissue biopsies consisting of cancer cells and stromal cells. Signatures obtained from fibroblasts upon serum stimulation [60], as well as growth factor derived signatures, such as a TGF-β gene expression signature in mouse hepatocytes [62], are well-described prognosticators in human breast cancer. In our study, we confirmed the validity and robustness of IGF-I derived signatures from primary breast and lung fibroblasts in four different human solid cancer datasets. Genes induced in primary breast fibroblasts upon IGF-I stimulation are predictive of outcome in breast cancer patients. N addition, the signature allows for the stratification of ER positive breast cancer patients into two groups with significantly different prognoses. Prognostication in this heterogeneous patient population is important for clinical decisions about adjuvant therapies in patients with ER positive breast cancer.

The ability to derive prognostic information from cancer stroma has already been shown by Finak *et al.* [63]. The gene expression signature of stromal cells obtained by laser capture microdissection (LCM), the stroma

derived prognostic predictor (SDPP), has been shown to be a prognostic marker in breast cancer. However, Finak *et al.* did not separate the different stromal components and, therefore, could not associate this signature to a specific cell type. In our study, we were able to specifically observe the effects of IGF-I on fibroblasts, which might be advantageous as targeted therapies are designed to specifically inhibit a signal at a particular cell type. Using laser capture microdissection, Roepman *et al.* managed to show that the genes expressed in the stroma are highly correlated with metastasis formation [64]. Specifically, they showed that 12% of the genes associated with lymph node metastasis in head-neck squamous cell carcinoma (HNSCC) are predominantly expressed in the stroma, 25% are tumour cell specific and the other genes are equally expressed in the tumour and the stroma. We speculate that the involvement of stroma-derived information might also be of importance in breast and in lung cancer. In our signatures, we found several of the genes that have been identified by Roepman *et al*. as being predominantly expressed in the stroma (ACTA1, TPM2, CDH2, COL5A1, COL5A2, HNRPL, TCF3). These segregated the patients into two groups with significantly different prognosis.

The IGF-I induced signatures in primary breast and in lung fibroblasts are similar to each other (Additional file 10) and to important, previously published signatures (Figure 4). The high reverse correlation of the IGF-I signature and the good-risk 70-genes signature supports the power of the IGF-I derived signature as a negative prognosticator in breast cancer. While the 'good-risk 70-genes signature' [49] was developed to predict freedom from metastasis in a top-down manner and validated in the same dataset of breast cancer patients from the NKI, the IGF-I induced signature is a marker for poor prognosis and is well connected to a defined *in vitro* biological system.

The IGF-I induced signature is also highly correlated to the wound signature [60], another strong prognostic signature in NKI dataset outperforming all known prognostic parameters so far. This is interesting, since a single growth factor, such as IGF-I, is able to induce a gene expression programme similar to the mix of undefined factors inherent in FBS. Using a fully defined stimulus in a concentration within the physiological range provides a simple and well-controlled *in vitro* model that enables specific experimental interventions to be made. Its effects can then be tested *in vivo*. Considering the notion by Sotirou [65] that proliferation is a main driver of the strong prognostic signatures such as the good-risk 70-genes signature and the wound signature facilitates speculation that IGF-I is one of the important factors responsible for the induction of proliferation. This does not exclude other, equally or more important, growth factors from inducing proliferation and up-regulation of proliferation associated genes.

We observed that both IGF-I signatures derived from lung and breast fibroblasts are exchangeable prognostic factors for the other cancer type, which allowed us to speculate that we could generalize this finding to other types of human solid cancer. The consistent response of fibroblasts (our data and [34,35,55]) to IGF-I might also help to explain the worse outcome of patients with elevated IGF-I levels in different cancer types [4-8], a finding that is not necessarily explained by the cancer cells themselves based on their IGF-receptor expression status on the cell surface. Specifically, since the correlation of the IGF-IR expression and patient outcome in human breast cancer is conflicting [66], the IGF-I induced gene expression signature showing the functional effects of IGF-I axis stimulation, which is correlated with the patients' clinical outcome, might be of interest when selecting patients who might benefit best from IGF-I blocking therapies.

IGF-I signalling is an emerging cancer drug target. *In vivo*, in mouse models, confirms that block IGF-I signalling demonstrate efficacy in inducing tumour regression and growth arrest [29] and sensitized cancer cells to conventional chemotherapeutic treatment and irradiation [67]. Exogenously added IGF binding protein I (IGFBP-1) inhibits IGF-I mediated growth of breast cancer cells [68,69]. Many other inhibitors of IGF signalling, applying different approaches [67], are currently under clinical investigation in phase I and II trials (reviewed in [29,70]). Some have already shown promising results, such as the phase II study on CP- 751, 851. This anti-insulin-like growth factor I receptor antibody, together with paclitaxel and carboplatin, was suggested to be safe and showed promising effectiveness in patients with non-small-cell lung cancer (NSCLC) showing the highest overall response rate of 78% in squamous cell carcinoma and 58% in adenocarcinomas [71]. Besides the monoclonal antibodies, there are small molecule inhibitors, such as XL228, that have blocking activity in the IGF1-R pathway and also in Src, fibroblast growth factor receptors (FGFR) and BCR-Abl pathways [72]. Although compounds that block IGF-I signalling demonstrate efficacy in inducing tumour regression and growth arrest *in vivo*, there is an emerging need to develop markers that predict a response to these therapies. We have tested the prognostic significance of our signature in patients with adenocarcinomas. In this group of patients, showing the lower response rate to IGF-I targeting therapies than squamous cell carcinomas [71], a better selection using a marker with predictive power would be especially beneficial. It might, therefore, be worthwhile to test whether or not the gene expression signatures

developed and described here are useful predictive markers for IGF-I signalling blockade.

Conclusions

The consistent and similar gene expression changes in human primary breast and lung fibroblasts suggest that the proliferative response to IGF-I is a general feature of stromal fibroblasts. Expression patterns of genes induced by IGF-I in primary breast and lung fibroblasts accurately predict outcomes in breast and in lung cancer patients. As IGF-I signalling is an emerging cancer drug target there is an emerging need to develop markers that predict a response to these therapies. Our IGF-I induced gene signatures derived from stromal fibroblasts might be promising predictors for the response to IGF-I targeted therapies.

Methods

Cell culture

Human primary fibroblasts CCL-171 and the human breast cancer cell line MCF-7 were obtained from American Type Culture Collection (ATTC, Atlanta, USA). Cells were propagated in Dulbecco's modified Eagle's medium (D-MEM, Invitrogen, Carlsbad, USA) supplemented with 10% heat inactivated FBS (Invitrogen), 4.5 g/lglucose, 4 mM L-glutamine and 100 U/ml penicillin and 100 μg/ml streptomycin (Gibco, Carlsbad, USA). Cells were maintained by regular passages when confluent. The study was approved by the Ethikkommission beider Basel, Switzerland (approval No. 271/05). Tumour and healthy tissue were obtained with consent from the patients who underwent surgery in University Hospital of Basel. For each patient, a sample of malignant tissue and a sample of side-matched healthy tissue were extracted by an experienced pathologist. The tissue was digested in a collagenase and RNase mix for 1 h and pressed through a 230 μm pore diameter sieve (Sigma Aldrich, St Louis, USA). The cells were cultured in a 1:1 v/v mixture of RPMI 1640 (Sigma Aldrich) and F12 Hamm (Gibco) medium supplemented with 12.5% FBS (Invitrogen), 2 mM Puryvat (Gibco), 4 mM L-glutamin (Gibco), 1 × Minimal Non-Essential Amino Acids (Gibco), 1 × RPMI 1640 Vitamins Solution (Sigma Aldrich), 100 U/ml penicillin and 100 μg/ml streptomycin (Gibco) and propagated until confluent. At this stage, cells were selected with anti-Fibroblast MicroBeads (Miltenyi Biotec, Gladbach, Germany) according to the manufacturer's instructions. All cells used in the experiments were kept in culture up to a maximum of 10 passages.

IGF-I stimulation

For the experiment, 30,000 cells/cm^2 were seeded in 3 mL of 5% FBS D-MEM for CCL-171 cells and 5% FBS RPMI 1640/F12 mix for primary cells for 6 h, so that they would attach. The cells were extensively washed with phosphate buffered saline and starved for 48 h in fresh low-serum medium (0.2% FBS), D-MEM and RPMI 1640/F12 mix for CCL-171 and primary cells, respectively. The cells were starved n order to reduce the effects of any stimulation from regular cell culture medium. The medium was subsequently replaced by fresh low-serum D-MEM with or without 50 ng/ml of IGF-I (human recombinant in *Escherichia coli*; Sigma Aldrich). The cells were stimulated for 24 h and the RNA was harvested to test the effects of IGF-I on mRNA expression patterns.

WST-1 proliferation assay

The proliferation reagent (Roche Diagnostics GmbH, Roche Applied Science, Basel, Switzerland) was used according to the manufacturer's instructions. In our setup, cells were plated in 96 well plates and starved for 48 h in low serum conditions. After, the cells were incubated in low-serum D-MEM with 50 ng/ml IGF-I over 24 h. n order to determine the cell numbers, the cells were stained with 10% WST-1 in low-serum D-MEM at 37°C, 5%CO_2 for 2 h. The absorbance was measured with an ELISA reader at a wavelength of 450 nm. The proliferation rate of IGF-I stimulated primary breast fibroblasts and CCL-171 cells was compared to a respective reference samples not stimulated with IGF-I.

RNA extraction and amplification

After aspirating the culture medium, the cell monolayer was washed once with phosphate buffered saline. The cells were lysed in a buffer containing guanidine isothiocyanate (RLT buffer, QIAGEN, CA, USA). The total RNA was extracted with the RNeasy kit (QIAGEN, CA, USA) according to the manufacturer's instructions. The RNA concentration was measured with a NanoDrop system spectrophotometer (ND-1000 Spectrophotometer Technologies, Wilmington, USA). The integrity of extracted RNA was checked by electrophoresis in a 1% agarose gel in MOPS buffer. For mRNA amplification, the Amino Allyl MasageAmp™ II aRNA Amplification Kit was used (Ambion, TX, USA). Amplification of mRNA out of 500 ng total RNA, the purification of cDNA, the *in vitro* transcription and the purification of aRNA were performed according to the manufacturer's instructions. Integrity and quantity of the amplified RNA was verified as described above.

Gene expression analysis using HEEBO microarrays

For global gene expression analysis, we used HEEBO. The HEEBO microarrays consist of 44,544 70mer probes, which include: (a) constitutive exonic probes (30,718); (b) alternatively spliced/skipped exonic probes (8,441); (c) non-coding RNA probes (196); (d) BCR/TCR genic/regional probes (372); (e) other probes (843); and (f) controls. HEEBO microarrays were produced at the Stanford Functional Genomic Facility (Stanford, USA).

Complete details regarding the clones on the arrays may be found at Stanford functional genomics facility website [73]. For microarray experiments, 8 μg amplified RNA (aRNA) were mixed with doping controls. Samples were vacuum dried, resolved in coupling buffer and labelled with Cy5 dye. Labelled samples were pooled with equal amounts of reverse coloured Cy3 labelled amplified reference RNA from Stratagene (Stratagene, CA, USA). The labelled aRNA was purified with AminoAllyl MasageAmp™ II aRNA Amplification Kit (Ambion) according tothe user manual and fragmented using fragmentation reagents (Ambion). The fragmented probe was added to a hybridization buffer containing Cot/PolyA/tRNA (0.05 μg/uL each), 0.3% SDS, 3.3 × SSC and supplemented with HEPES buffer. Following a denaturing step at 100° C, the probe was placed on the microarray for competitive hybridization. After 18 h, slides with hybridized probes were sequentially washed and immediately dried in an ozone free environment and scanned using an Axon Scanner 4100A (Axon Instruments, CA, USA). The gene expression profiles of primary fibroblasts, together with accompanying clinical data are available on SMD database papers' webpage [39]. In addition, the raw data have been deposited in NCBI's Gene Expression Omnibus [74] and are accessible through GEO Series accession number GSE18955 http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE18955.

Data analysis and clustering

Microarray fluorescent image analysis was performed using the software Genepix Pro version 5.0 3.0.6.89 (Axon Instruments). Spots with obvious array artifacts or poor technical quality were manually removed from any further analysis. Raw data files were stored in the Stanford Microarray Database [39]. The data used for the paper are available at the accompanying website at Stanford Microarray Database [39]. Data were expressed as the $\log^2$ ratio of fluorescence intensities of the sample and the reference for each element on the array. A sequential data filtering procedure was applied to include only measurements fulfilling our quality requirements (data with regression correlation bigger than 0.6 and Cy3 channel or Cy5 channel mean intensity over median background intensity bigger than 1.5). Genes that did not meet these criteria for at least 60% of the measurements across the experimental samples were excluded from further analysis. We rejected elements that did not have at least a 1.5-fold deviation from the mean in at least two samples. Data were evaluated by unsupervised hierarchical clustering [75] and displayed using Treeview software [41]. For the stimulation experiments, in order to emphasize the effect of IGF-I treatment, the results for each gene were standardized for each gene individually to the non-treated samples. In order to standardize them, we subtracted an average value of non-treated samples from each gene expression value for each cell type separately. This was performed in order to highlight those genes whose expression level changed upon treatment. Extraction of fibroblast gene signature and differentially expressed clusters was based on the correlation within the cluster nodes and, therefore, not randomly selected or based on an arbitrary cut- off.

In order to compare the gene expression profile of CCL-171 and MCF-7 cells in response to IGF-I, we merged the filtered, standardized gene expression profiles of both cell lines. We then manually excluded samples with a high standard deviation between the biological replicates and those missing gene expression data. Gene expression data for different clones representing one gene were averaged. A set of 566 unique genes was hierarchically clustered in an unsupervised manner [75] and displayed using Treeview software [41].

SAM

For primary fibroblasts, two-class SAM was applied [43]. One class was formed by normal and carcinoma associated fibroblasts starved in low serum medium and the other by the same cells treated with IGF-I. In order to increase the sensitivity, we paired our samples.

Human cancer datasets

A dataset containing gene expression patterns from advanced breast cancers was previously described by Sorlie *et al.* as Norway/Stanford dataset [45,46]. Expression measurements for each gene and array were mean centred. The list of 208 unique genes building breast fibroblast derived IGF-I signature was extracted from the Norway/Stanford dataset. In order to overcome possible overweighting of clones from Unigene clusters that were matched to more than one probe on the Sorlie array, expression values derived from probes matched to the same Unigene cluster were averaged. Only genes that had >80% data values present and tumour samples from patients having complete clinical data were used. The resulting dataset was subjected to average linkage hierarchical clustering [75] and displayed with Treeview [41].

Disease specific survival analysis was based on death from the disease and patients were censored at the last follow up. Patients who died from other causes were considered alive and not censored. Kaplan-Meier survival curves were compared using R package survival fitting a Cox proportional hazards regression model [76].

The dataset for early stage breast cancer contained 295 breast cancer specimens analysed on a 25,000 spot oligonucleotide array, as described previously [44]. In brief, patients were diagnosed and treated at the Netherlands Cancer Institute (NKI) for early stage breast cancer (stage I and II) between 1984 and 1995. The clinical data was updated in January 2005. The median follow-

up for patients still alive is 12.3 years. Expression data from the NKI dataset were extracted as described above for the Norway/Stanford dataset. Distant metastases were analysed as a first event only (distant metastasis-free probability). Any patient who developed a local recurrence, axillary recurrence, contralateral breast cancer or a second primary cancer (except for non-melanoma skin cancer), was censored at that time and subsequent distant metastases were not analysed. This is based on the theoretical possibility that the locally recurrent or second primary cancers could be a source for distant metastases. An ipsilateral supra-clavicular recurrence was soon followed by a distant metastasis in all but one patient. Thus, an ipsilateral supra-clavicular recurrence was considered the first clinical evidence for metastatic disease for this analysis and patients were not censored at the time of ipsilateral supra-clavicular recurrence. Overall survival was analysed based on death from any cause and patients were censored at last follow up. Kaplan-Meier survival curves were fitted using a Cox proportional hazards regression model (R survival package) [76].

The dataset published by Garber and colleagues [50] contains global gene expression profiles for 67 human lung cancers derived from 56 patients with survival data for 24 patients. The dataset published by Bhattacharjee [51] contains mRNA expression levels of 12,600 transcript sequences in 186 lung tumor samples, including 139 adenocarcinomas resected from the lung. Of these, 125 samples were associated with clinical data (some patients in multiple runs). The Bhattacharjee dataset was obtained from the Broad Institute website [77] and Garber dataset from SMD publication webpage [39]. The list of 370 unique genes building fibroblast derived IGF-I signature was extracted from the Garber and Bhatacharjee datasets as described above for breast cancer datasets. Equally, the resulting dataset was subjected to average linkage hierarchical clustering [75] and displayed with Treeview [41]. Overall survival was analysed based on death from any cause and patients were censored at last follow up. Disease specific survival analysis was based on death from the disease and patients were censored at last follow up. Patients who died from other causes were considered alive and not censored. Kaplan-Meier survival curves were fitted using a Cox proportional hazards regression model (R package 'survival') [76].

Centroid correlation

The method of calculating the centroid for each patient was previously described by Sorlie [45]. Briefly, the centroids for genes representing breast fibroblast derived IGF-I signature and fibroblast derived IGF-I signature, as well as other signatures, were calculated based on the NKI dataset. To test for similarities between the signatures, we checked the correlation between values of different centroids for one patient. The correlation was calculated using Pearson correlation coefficient with R software [76].

Continuous scoring

The stratification of patients within the NKI and Bhattacharjee datasets was conducted according to the previously described methodology [60,61] based on a continuous score derived from the signatures. Briefly, the average expression level of each signature was calculated for each patient attributing a score. The patients were then divided into two groups separating them by the median value of the continuous scores. Kaplan-Meier survival curves for the two groups were plotted and the statistical significance was determined using a Cox proportional hazards model (R package 'survival') [76].

GO::TermFinder analysis

GO::TermFinder takes a list of genes as input, and determines whether those genes have any gene ontology (GO) terms overrepresented in their combined set of annotations compared to what would be expected by chance from a randomly selected group of genes from the background population of all genes [39,40]. In our analysis, we used the full gene lists from parental heat maps as a file to calculate the frequency of particular annotations in a background file and the gene lists from specific clusters coming from same heat map to calculate the frequency of particular annotations in the defined cluster. For a SAM-derived signature, we used a gene list that was an input file for SAM analysis.

General statistic methods

Normally distributed data were compared using a Student's *t*-test. When the multiple comparisons were necessary, the data were analysed with ANOVA. Differences were considered as statistically significant when $P < 0.05$. *T*-tests and ANOVA analysis were done using R software (R package 'stats') [76].

Additional file 1: Table S1. List of genes building the fibroblast derived insulin-like growth factor-1 (IGF-I) signature.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1741-7015-8-1-S1.PDF]

Additional file 2: Figure S2. Graphical visualization of the output from GO::Termfinder for biological process ontology. GOgraph layout that includes the significant GO nodes up-regulated in CCL-171 cells, derived from 325 clones compared to a background of 2133 clones. The colour of the nodes is an indication of their Bonferroni corrected *P*-value (orange <= 1e-10; yellow 1e-10 to 1e-8; green 1e-8 to 1e-6; cyan 1e-6 to 1e-4; blue 1e-4 to 1e-2; tan > 0.01).
Click here for file
[http://www.biomedcentral.com/content/supplementary/1741-7015-8-1-S2.TIFF]

Additional file 3: Figure S1. Distinct default gene expression profiles of human lung fibroblasts and breast tumour cells. Genes are presented in rows and experiments in columns. Both cell types demonstrate a clearly

distinct default gene expression profile, typical for epithelial and mesenchymal cells. Gene markers typical for mesenchymal (FN1, CDH2, VIM) and epithelial/tumour cells (CDH1, TPD52, BMP-7) are marked. Additionally, examples of proliferation associated genes up-regulated in MCF-7 cells by default are shown.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1741-7015-8-1-S3.TIFF]

Additional file 4: Figure S3. Box-and-whisker plot illustrating the average expression level of fibronectin (FN1), N-cadherin (CDH2) and E-cadherin (CDH1) in primary fibroblasts. Insulin-like growth factor (IGF-I) does not affect the expression level of mesenchymal and epithelial markers in primary breast fibroblasts (data not shown).
Click here for file
[http://www.biomedcentral.com/content/supplementary/1741-7015-8-1-S4.TIFF]

Additional file 5: Table S2. List of genes up-regulated (the breast fibroblast derived insulin-like growth factor-1 [IGF-I] signature) and down-regulated in primary breast fibroblasts upon IGF-I stimulation.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1741-7015-8-1-S5.PDF]

Additional file 6: Figure S4. Graphical visualization of the output from GO::Termfinder for biological process ontology. GOgraph layout that includes the significant GO nodes up-regulated in primary breast fibroblasts, derived from 186 clones compared to a background of 8918 clones. The colour of the nodes is an indication of their Bonferroni corrected *P*-value (orange <= 1e-10; yellow 1e-10 to 1e-8; green 1e-8 to 1e-6; cyan 1e-6 to 1e-4; blue 1e-4 to 1e-2; tan > 0.01).
Click here for file
[http://www.biomedcentral.com/content/supplementary/1741-7015-8-1-S6.TIFF]

Additional file 7: Figure S6. Relationship of expression level of breast fibroblast derived insulin-like growth factor-1 (IGF-I) signature with distant metastasis free and overall survival applying continuous scoring. A. Continuous score based on average expression level of the signature in Netherlands Cancer Institute (NKI) patients. Colours correspond to score below (yellow) or above (blue) the median (red line). Overall (B) and metastasis free survival (C) analysis using a continuous score resulting from breast fibroblast derived IGF-I signature in early stage breast cancer patients from the NKI.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1741-7015-8-1-S7.PDF]

Additional file 8: Table S3. The detailed list of correlation values of breast fibroblast derived insulin-like growth factor-1 (IGF-I) signature to the previously published signatures and fibroblast derived IGF-I signature.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1741-7015-8-1-S8.PDF]

Additional file 9: Figure S7. Relationship of expression level of breast fibroblast derived insulin-like growth factor-1 (IGF-I) signature with overall survival and disease specific survival applying continuous scoring. A. Continuous score based on average expression level of the signature in Bhattacharjee dataset patients. Colours correspond to score below (yellow) or above (blue) the median (red line). Overall (B) and disease specific survival (C) analysis using a continuous score resulting from breast fibroblast derived IGF-I signature in Bhattacharjee dataset patients.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1741-7015-8-1-S9.PDF]

Additional file 10: Figure S5. Correlation of the fibroblast derived insulin-like growth factor-1 (IGF-I) signature and the breast fibroblast IGF-I induced signature centroids in the Netherlands Cancer Institute dataset. Pearson correlations for the signature and the *P* value are shown in the lower right part of the plot.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1741-7015-8-1-S10.TIFF]

Acknowledgements
We would like to thank Michael Fero, Elena Seraia and the staff of the Stanford Functional Genomics Facility for supplying us with the human cDNA microarrays that were used for this study. We also thank Kathy Ball, Janos Demeter and the staff from SMD for their support with the microarray data storage and analysis tools. We are grateful to Christian Ludwig for releasing MB from part of his clinical duties to support this project and Florent Baty for his input concerning the R-software for the statistical analyses. MB is a fellow of the Schoenmakers-Müller Foundation, Basel, Switzerland. This work was supported in part by the Swiss National Science Foundation grant 320000-112794/1 and a grant from Oncosuisse.

Author details
[1]Department of Biomedicine, University of Basel, Hebelstrasse 20, CH-4031 Basel, Switzerland. [2]Division of Medical Oncology, Department of Internal Medicine, St Claraspital, Kleinriehenstrasse 20, CH-4016 Basel, Switzerland. [3]Division of Medical Oncology, Department of Internal Medicine, University Hospital Basel, Petersgraben 4, CH-4031 Basel, Switzerland. [4]Department of Gynecology and Obstetrics, University Hospital Basel, Petersgraben 4, CH-4031 Basel, Switzerland.

Authors' contributions
MR and MB designed the research, analysed and interpreted the data. MR and BV performed the research. MB, RZ recruited the patients and analysed clinical data. All authors have been involved in drafting the manuscript and revising it critically for important intellectual content and have given final approval of the version to be published.

Competing interests
The authors declare that they have no competing interests.

Received: 9 December 2009
Accepted: 5 January 2010 Published: 5 January 2010

References

1. Baserga R, Peruzzi F, Reiss K: **The IGF-1 receptor in cancer biology.** *Int J Cancer* 2003, **107**:873-877.
2. Jones JI, Clemmons DR: **Insulin-like growth factors and their binding proteins: biological actions.** *Endocr Rev* 1995, **16**:3-34.
3. Marshman E, Streuli CH: **Insulin-like growth factors and insulin-like growth factor binding proteins in mammary gland function.** *Breast Cancer Res* 2002, **4**:231-239.
4. Bohlke K, Cramer DW, Trichopoulos D, Mantzoros CS: **Insulin-like growth factor-I in relation to premenopausal ductal carcinoma *in situ* of the breast.** *Epidemiology* 1998, **9**:570-573.
5. Chan JM, Stampfer MJ, Giovannucci E, Gann PH, Ma J, Wilkinson P, Hennekens CH, Pollak M: **Plasma insulin-like growth factor-I and prostate cancer risk: a prospective study.** *Science* 1998, **279**:563-566.
6. Hankinson SE, Willett WC, Colditz GA, Hunter DJ, Michaud DS, Deroo B, Rosner B, Speizer FE, Pollak M: **Circulating concentrations of insulin-like growth factor-I and risk of breast cancer.** *Lancet* 1998, **351**:1393-1396.
7. Ma J, Pollak MN, Giovannucci E, Chan JM, Tao Y, Hennekens CH, Stampfer MJ: **Prospective study of colorectal cancer risk in men and plasma levels of insulin-like growth factor (IGF)-I and IGF-binding protein-3.** *J Natl Cancer Inst* 1999, **91**:620-625.
8. Yu H, Spitz MR, Mistry J, Gu J, Hong WK, Wu X: **Plasma levels of insulin-like growth factor-I and lung cancer risk: a case-control analysis.** *J Natl Cancer Inst* 1999, **91**:151-156.
9. Lann D, LeRoith D: **The role of endocrine insulin-like growth factor-I and insulin in breast cancer.** *J Mammary Gland Biol Neoplasia* 2008, **13**:371-379.
10. Dimitriadis G, Parry-Billings M, Bevan S, Dunger D, Piva T, Krause U, Wegener G, Newsholme EA: **Effects of insulin-like growth factor I on the**

rates of glucose transport and utilization in rat skeletal muscle *in vitro*. *Biochem J* 1992, **285(Pt 1)**:269-274.
11. Le Roith D, Bondy C, Yakar S, Liu JL, Butler A: **The somatomedin hypothesis: 2001.** *Endocr Rev* 2001, **22**:53-74.
12. Farrelly N, Lee YJ, Oliver J, Dive C, Streuli CH: **Extracellular matrix regulates apoptosis in mammary epithelium through a control on insulin signaling.** *J Cell Biol* 1999, **144**:1337-1348.
13. Pollak MN, Schernhammer ES, Hankinson SE: **Insulin-like growth factors and neoplasia.** *Nat Rev Cancer* 2004, **4**:505-518.
14. Stull MA, Richert MM, Loladze AV, Wood TL: **Requirement for IGF-I in epidermal growth factor-mediated cell cycle progression of mammary epithelial cells.** *Endocrinology* 2002, **143**:1872-1879.
15. Deeks S, Richards J, Nandi S: **Maintenance of normal rat mammary epithelial cells by insulin and insulin-like growth factor 1.** *Exp Cell Res* 1988, **174**:448-460.
16. Shamay A, Cohen N, Niwa M, Gertler A: **Effect of insulin-like growth factor I on deoxyribonucleic acid synthesis and galactopoiesis in bovine undifferentiated and lactating mammary tissue *in vitro*.** *Endocrinology* 1988, **123**:804-809.
17. Richert MM, Wood TL: **The insulin-like growth factors (IGF) and IGF type I receptor during postnatal growth of the murine mammary gland: sites of messenger ribonucleic acid expression and potential functions.** *Endocrinology* 1999, **140**:454-461.
18. Neuenschwander S, Schwartz A, Wood TL, Roberts CT Jr, Hennighausen L, LeRoith D: **Involution of the lactating mammary gland is inhibited by the IGF system in a transgenic mouse model.** *J Clin Invest* 1996, **97**:2225-2232.
19. Ruan W, Catanese V, Wieczorek R, Feldman M, Kleinberg DL: **Estradiol enhances the stimulatory effect of insulin-like growth factor-I (IGF-I) on mammary development and growth hormone-induced IGF-I messenger ribonucleic acid.** *Endocrinology* 1995, **136**:1296-1302.
20. Dunn SE, Hardman RA, Kari FW, Barrett JC: **Insulin-like growth factor 1 (IGF-1) alters drug sensitivity of HBL100 human breast cancer cells by inhibition of apoptosis induced by diverse anticancer drugs.** *Cancer Res* 1997, **57**:2687-2693.
21. Gooch JL, Berg Van Den CL, Yee D: **Insulin-like growth factor (IGF)-I rescues breast cancer cells from chemotherapy-induced cell death–proliferative and anti-apoptotic effects.** *Breast Cancer Res Treat* 1999, **56**:1-10.
22. Papa V, Gliozzo B, Clark GM, McGuire WL, Moore D, Fujita-Yamaguchi Y, Vigneri R, Goldfine ID, Pezzino V: **Insulin-like growth factor-I receptors are overexpressed and predict a low risk in human breast cancer.** *Cancer Res* 1993, **53**:3736-3740.
23. Resnik JL, Reichart DB, Huey K, Webster NJ, Seely BL: **Elevated insulin-like growth factor I receptor autophosphorylation and kinase activity in human breast cancer.** *Cancer Res* 1998, **58**:1159-1164.
24. Railo MJ, von Smitten K, Pekonen F: **The prognostic value of insulin-like growth factor-I in breast cancer patients. Results of a follow-up study on 126 patients.** *Eur J Cancer* 1994, **30A**:307-311.
25. Resnicoff M, Coppola D, Sell C, Rubin R, Ferrone S, Baserga R: **Growth inhibition of human melanoma cells in nude mice by antisense strategies to the type 1 insulin-like growth factor receptor.** *Cancer Res* 1994, **54**:4848-4850.
26. Valentinis B, Morrione A, Taylor SJ, Baserga R: **Insulin-like growth factor I receptor signaling in transformation by src oncogenes.** *Mol Cell Biol* 1997, **17**:744-3754.
27. Dunn SE, Ehrlich M, Sharp NJ, Reiss K, Solomon G, Hawkins R, Baserga R, Barrett JC: **A dominant negative mutant of the insulin-like growth factor-I receptor inhibits the adhesion, invasion, and metastasis of breast cancer.** *Cancer Res* 1998, **58**:3353-3361.
28. Nagle JA, Ma Z, Byrne MA, White MF, Shaw LM: **Involvement of insulin receptor substrate 2 in mammary tumor metastasis.** *Mol Cell Biol* 2004, **24**:9726-9735.
29. Hartog H, Wesseling J, Boezen HM, Graaf van der WT: **The insulin-like growth factor 1 receptor in cancer: old focus, new future.** *Eur J Cancer* 2007, **43**:1895-1904.
30. Sachdev D, Yee D: **The IGF system and breast cancer.** *Endocr Relat Cancer* 2001, **8**:197-209.
31. Yee D, Paik S, Lebovic GS, Marcus RR, Favoni RE, Cullen KJ, Lippman ME, Rosen N: **Analysis of insulin-like growth factor I gene expression in malignancy: evidence for a paracrine role in human breast cancer.** *Mol Endocrinol* 1989, **3**:509-517.
32. Pekonen F, Partanen S, Makinen T, Rutanen EM: **Receptors for epidermal growth factor and insulin-like growth factor I and their relation to steroid receptors in human breast cancer.** *Cancer Res* 1988, **48**:1343-1347.
33. Surmacz E: **Function of the IGF-I Receptor in Breast Cancer.** *J Mammary Gland Biology Neoplasia* 2000, **5**:95-105.
34. Scarpa RC, Carraway RE, Cochrane DE: **Insulin-like growth factor (IGF) induced proliferation of human lung fibroblasts is enhanced by neurotensin.** *Peptides* 2005, **26**:2201-2210.
35. Yateman ME, Claffey DC, Cwyfan Hughes SC, Frost VJ, Wass JA, Holly JM: **Cytokines modulate the sensitivity of human fibroblasts to stimulation with insulin-like growth factor-I (IGF-I) by altering endogenous IGF-binding protein production.** *J Endocrinol* 1993, **137**:151-159.
36. Johansson GS, Chisalita SI, Arnqvist HJ: **Human microvascular endothelial cells are sensitive to IGF-I but resistant to insulin at the receptor level.** *Mol Cell Endocrinol* 2008, **296**:8-63.
37. Bendall SC, Stewart MH, Menendez P, George D, Vijayaragavan K, Werbowetski-Ogilvie T, Ramos-Mejia V, Rouleau A, Yang J, Bossé M, Lajoie G, Bhatia M: **IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells *in vitro*.** *Nature* 2007, **448**:1015-1021.
38. Perou CM, Sørlie T, Eisen MB, Rijn van de M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale AL, Brown PO, Botstein D: **Molecular portraits of human breast tumours.** *Nature* 2000, **406**:747-752.
39. **Stanford Miocroarray Database.** http://smd.stanford.edu/index.shtml.
40. Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, Sherlock G: **GO::TermFinder-open source software for accessing gene ontology information and finding significantly enriched gene ontology terms associated with a list of genes.** *Bioinformatics* 2004, **20**:3710-3715.
41. Page RD: **TreeView: an application to display phylogenetic trees on personal computers.** *Comput Appl Biosci* 1996, **12**:357-358.
42. Rinn JL, Bondre C, Gladstone HB, Brown PO, Chang HY: **Anatomic demarcation by positional variation in fibroblast gene expression programs.** *PLoS Genet* 2006, **2**:e119.
43. Tusher VG, Tibshirani R, Chu G: **Significance analysis of microarrays applied to the ionizing radiation response.** *Proc Natl Acad Sci USA* 2001, **98**:5116-5121.
44. Vijver van de MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, Velde van der T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R: **A gene-expression signature as a predictor of survival in breast cancer.** *N Engl J Med* 2002, **347**:1999-2009.
45. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lønning PE, Brown PO, Børresen-Dale AL, Botstein D: **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proc Natl Acad Sci USA* 2003, **100**:8418-8423.
46. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, Rijn van de M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lønning P, Børresen-Dale AL: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci USA* 2001, **98**:10869-10874.
47. Kim DW, Lee KY, Lee KH, Lee D: **Towards clustering of incomplete microarray data without the use of imputation.** *Bioinformatics* 2007, **23**:107-113.
48. Chang HY, Nuyten DS, Sneddon JB, Hastie T, Tibshirani R, Sørlie T, Dai H, He YD, van't Veer LJ, Bartelink H, Rijn van de M, Brown PO, Vijver van de MJ: **Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival.** *Proc Natl Acad Sci USA* 2005, **102**:3738-3743.
49. van't Veer LJ, Dai H, Vijver van de MJ, He YD, Hart AA, Mao M, Peterse HL, Kooy van der K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH: **Gene expression profiling predicts clinical outcome of breast cancer.** *Nature* 2002, **415**:530-536.
50. Garber ME, Troyanskaya OG, Schluens K, Petersen S, Thaesler Z, Pacyna-Gengelbach M, Rijn van de M, Rosen GD, Perou CM, Whyte RI, Altman RB, Brown PO, Botstein D, Petersen I: **Diversity of gene expression in adenocarcinoma of the lung.** *Proc Natl Acad Sci USA* 2001, **98**:13784-13789.
51. Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, Vasa P, Ladd C, Beheshti J, Bueno R, Gillette M, Loda M, Weber G, Mark EJ, Lander ES, Wong W, Johnson BE, Golub TR, Sugarbaker DJ, Meyerson M: **Classification**

of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci USA* 2001, **98**:13790-13795.
52. Pacher M, Seewald MJ, Mikula M, Oehler S, Mogg M, Vinatzer U, Eger A, Schweifer N, Varecka R, Sommergruber W, Mikulits W, Schreiber M: **Impact of constitutive IGF1/IGF2 stimulation on the transcriptional program of human breast cancer cells.** *Carcinogenesis* 2007, **28**:49-59.
53. Creighton CJ, Casa A, Lazard Z, Huang S, Tsimelzon A, Hilsenbeck SG, Osborne CK, Lee AV: **Insulin-like growth factor-I activates gene transcription programs strongly associated with poor breast cancer prognosis.** *J Clin Oncol* 2008, **26**:4078-4085.
54. Paik S: **Expression of IGF-I and IGF-II mRNA in breast tissue.** *Breast Cancer Res Treatment* 1992, **22**:31-38.
55. Dupont J, Khan J, Qu BH, Metzler P, Helman L, LeRoith D: **Insulin and IGF-1 induce different patterns of gene expression in mouse fibroblast NIH-3T3 cells: identification by cDNA microarray analysis.** *Endocrinology* 2001, **142**:4969-4975.
56. Sasaki H, Yu CY, Dai M, Tam C, Loda M, Auclair D, Chen LB, Elias A: **Elevated serum periostin levels in patients with bone metastases from breast but not lung cancer.** *Breast Cancer Res Treat* 2003, **77**:245-252.
57. Shao R, Bao S, Bai X, Blanchette C, Anderson RM, Dang T, Gishizky ML, Marks JR, Wang XF: **Acquired expression of periostin by human breast cancers promotes tumor angiogenesis through up-regulation of vascular endothelial growth factor receptor 2 expression.** *Mol Cell Biol* 2004, **24**:3992-4003.
58. Ruiz C, Huang W, Hegi ME, Lange K, Hamou MF, Fluri E, Oakeley EJ, Chiquet-Ehrismann R, Orend G: **Growth promoting signaling by tenascin-C [corrected].** *Cancer Res* 2004, **64**:7377-7385.
59. Akiri G, Sabo E, Dafni H, Vadasz Z, Kartvelishvily Y, Gan N, Kessler O, Cohen T, Resnick M, Neeman M, Neufeld G: **Lysyl oxidase-related protein-1 promotes tumor fibrosis and tumor progression *in vivo*.** *Cancer Res* 2003, **63**:1657-1666.
60. Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K, Chi JT, Rijn van de M, Botstein D, Brown PO: **Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds.** *PLoS Biol* 2004, **2**:E7.
61. Chi JT, Wang Z, Nuyten DS, Rodriguez EH, Schaner ME, Salim A, Wang Y, Kristensen GB, Helland A, Børresen-Dale AL, Giaccia A, Longaker MT, Hastie T, Yang GP, Vijver van de MJ, Brown PO: **Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers.** *PLoS Med* 2006, **3**:e47.
62. Coulouarn C, Factor VM, Thorgeirsson SS: **Transforming growth factor-beta gene expression signature in mouse hepatocytes predicts clinical outcome in human cancer.** *Hepatology* 2008, **47**:2059-2067.
63. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, Chen H, Omeroglu G, Meterissian S, Omeroglu A, Hallett M, Park M: **Stromal gene expression predicts clinical outcome in breast cancer.** *Nat Med* 2008, **14**:518-527.
64. Roepman P, de Koning E, van Leenen D, de Weger R, Kummer JA, Slootweg P, Holstege F: **Dissection of a metastatic gene expression signature into distinct components.** *Genome Biology* 2006, **7**:R117.
65. Sotiriou C, Pusztai L: **Gene-expression signatures in breast cancer.** *N Engl J Med* 2009, **360**:790-800.
66. Sachdev D: **Regulation of breast cancer metastasis by IGF signaling.** *J Mammary Gland Biology Neoplasia* 2008, **13**:431-441.
67. Bahr C, Groner B: **The insulin like growth factor-1 receptor (IGF-1R) as a drug target: novel approaches to cancer therapy.** *Growth Horm IGF Res* 2004, **14**:287-295.
68. McGuire WL Jr, Jackson JG, Figueroa JA, Shimasaki S, Powell DR, Yee D: **Regulation of insulin-like growth factor-binding protein (IGFBP) expression by breast cancer cells: use of IGFBP-1 as an inhibitor of insulin-like growth factor action.** *J Natl Cancer Inst* 1992, **84**:1336-1341.
69. Figueroa JA, Sharma J, Jackson JG, McDermott MJ, Hilsenbeck SG, Yee D: **Recombinant insulin-like growth factor binding protein-1 inhibits IGF-I, serum, and estrogen-dependent growth of MCF-7 human breast cancer cells.** *J Cell Physiol* 1993, **157**:229-236.
70. Hewish M, Chau I, Cunningham D: **Insulin-like growth factor 1 receptor targeted therapeutics: novel compounds and novel treatment strategies for cancer medicine.** *Recent Pat Anticancer Drug Discov* 2009, **4**:54-72.
71. Karp DD, Paz-Ares LG, Novello S, Haluska P, Garland L, Cardenal F, Blakely LJ, Eisenberg PD, Langer CJ, Blumenschein G Jr, Johnson FM, Green S, Gualberto A: **Phase II study of the anti-insulin-like growth factor type 1 receptor antibody CP-751,871 in combination with paclitaxel and carboplatin in previously untreated, locally advanced, or metastatic non-small-cell lung cancer.** *J Clin Oncol* 2009, **27**:2516-2522.
72. Smith DC, Britten C, Clary DO, Nguyen LT, Woodard P, Hurwitz HI: **A phase I study of XL228, a potent IGF1R/AURORA/SRC inhibitor, in patients with solid tumors or hematologic malignancies.** *J Clin Oncol (Meeting Abstracts)* 2009, **27**.
73. **Stanford Functional Genomics Facility.** http://www.microarray.org/.
74. Edgar R, Domrachev M, Lash AE: **Gene expression omnibus: NCBI gene expression and hybridization array data repository.** *Nucleic Acids Res* 2002, **30**:207-210.
75. Eisen MB, Spellman PT, Brown PO, Botstein D: **Cluster analysis and display of genome-wide expression patterns.** *Proc Natl Acad Sci USA* 1998, **95**:14863-14868.
76. R_Development_Core_Team: *R: A Language and Environment for Statistical Computing* Vienna: R Foundation for Statistical Computing 2006.
77. **Classification of Human Lung Carcinomas by mRNA Expression Profiling Reveals Distinct Adenocarcinoma Sub-classes.** 2001http://www.broad.mit.edu/mpr/lung/.

Pre-publication history
The pre-publication history for this paper can be accessed here:http://www.biomedcentral.com/1741-7015/8/1/prepub

doi:10.1186/1741-7015-8-1
Cite this article as: Rajski *et al.*: IGF-I induced genes in stromal fibroblasts predict the clinical outcome of breast and lung cancer patients. *BMC Medicine* 2010 **8**:1.

Additional files provided with this submission:

Additional file 1: sup1.pdf, 51K
http://www.biomedcentral.com/imedia/9152287113377051/supp1.pdf
Additional file 2: sup2.tiff, 164K
http://www.biomedcentral.com/imedia/1589079261337705/supp2.tiff
Additional file 3: sup3.tiff, 107K
http://www.biomedcentral.com/imedia/7516222843377058/supp3.tiff
Additional file 4: sup4.tiff, 197K
http://www.biomedcentral.com/imedia/4810054833377052/supp4.tiff
Additional file 5: sup5.pdf, 116K
http://www.biomedcentral.com/imedia/1680310293377052/supp5.pdf
Additional file 6: suppl6.tiff, 100K
http://www.biomedcentral.com/imedia/2045310524338562/supp6.tiff
Additional file 7: suppl7.pdf, 90K
http://www.biomedcentral.com/imedia/1805348435338561/supp7.pdf
Additional file 8: suppl8.pdf, 28K
http://www.biomedcentral.com/imedia/2339397633856135/supp8.pdf
Additional file 9: suppl9.pdf, 87K
http://www.biomedcentral.com/imedia/1482681807338561/supp9.pdf
Additional file 10: suppl10.tiff, 509K
http://www.biomedcentral.com/imedia/1537671566338561/supp10.tiff

Conclusions and perspectives

The results section deals with aspects of two biological phenomena having an impact on breast cancer; mutual interaction of breast cancer cells with endothelial cells and the influence of a growth factor on primary normal and tumor associated breast fibroblasts. The changes induced due to both aspects that we explored, might have an impact on the course of breast cancer. The observed gene expression changes allowed the formulation of gene expression signatures, which were evaluated within publically available gene expression databases, to validate the relevance of *in vitro* results *in vivo*.

Fibroblast derived IGF–I signatures

We showed that primary normal breast fibroblasts, carcinoma-associated breast fibroblasts and fibroblasts from the lung are sensitive to IGF–I stimulation, and that the fibroblasts of the different origins show a specific gene expression and phenotypical response to this stimulus. Moreover, genes up-regulated in primary breast and lung fibroblasts upon IGF-I stimulation have a prognostic significance in human breast cancer and lung adenocarcinomas. Gene expression patterns induced by IGF-I in primary breast fibroblasts have several advantages compared to the previously established “wound signature” [117] which is triggered by serum consisting of an unspecific mixture of potent stimuli on a mix of fibroblasts from different sites of the body. First, the gene expression pattern presented here was derived from a strictly controlled *in vitro* system with a defined concentration of recombinant IGF-I. Second, since it is known that carcinoma-associated fibroblasts vary from normal fibroblasts [28-29, 132] and that fibroblasts from different body sites have different default gene expression programs [134-135], site-specific breast primary fibroblasts were used to examine the subtle changes characteristic for these cells only. The signature from IGF-I stimulated primary breast fibroblasts was able to group breast cancer patients with similar clinical outcome together. Similarly, the signature derived from primary lung fibroblasts enabled the separation of lung cancer patients with distinct survival rates. Furthermore, the power of both signatures, derived from

breast and lung fibroblasts, to cross-predict patient clinical performance in both types of solid tumors might suggest that gene expression programs represented in IGF-I stimulated fibroblasts is a general feature of solid tumors. In contrast to the “70-genes signature” [104] our gene expression patterns were derived from an unsupervised analysis. This approach assesses the correlation between the induced gene expression profile and clinical data in an independent way. Since, the gene expression signatures stratify the patients into groups with different overall survival, we speculate that the represented features are important for breast and lung cancer biology. Finak *et al.* [120] compared genes expressed in normal stroma to genes specifically induced in the whole LCM dissected tumor associated stroma. In contrast to Finak, the gene expression IGF-I-induced signatures are derived from a single cell type. As already suggested in the introduction, this gives an insight into actions taking place in the specific cell fraction. For example numerous soluble factors (POSTN, TNC, LOXL1) are induced in primary breast fibroblasts upon IGF-I stimulation. POSTN, TNC and a protein similar to LOXL1 were previously reported to be involved in tumor progression [140-143]. Moreover, the IGF-I-induced signatures from primary breast and lung fibroblasts are similar to each other and to other, significant, previously published signatures, like “wound” and “70-genes” and gene expression signature of “luminal B” group of breast cancer [88, 104, 117]. These signatures represent important aspects of tumor biology what imposes the role of processes induced by IGF-I in tumor stroma in cancer progression. The additional comparison of IGF-I actions on tumor cells and fibroblasts showed that these cell types have different changes in their gene expression patterns in response to the stimulus and therefore may behave differently. The IGF-I signaling pathway is important in the development of many tissues [144] and is closely related to insulin receptor signaling [145]. Both processes, development and insulin dependent metabolism, are vital in mammary gland homeostasis. Since IGF-I targeting affects both tumor and carcinoma associated fibroblasts, it was speculated that knowledge about the fibroblasts-characteristic response might be of importance to avoid targeting of these vital functions in non-

carcinoma associated tissues [146]. Publicly available gene expression profiles (GEO: GSE3398) might help in future to address these issues involved in development of IGF-I blocking therapies. It can also be hypothesized that it may be of clinical significance to analyze the presented signatures as potential response predictors for these therapies.

Tumor endothelium interaction derived signature

The second part of the thesis is focused on the gene expression changes induced due to tumor-endothelial cell interaction. This work is an extension of experiments involving *ex-vivo* co-culture models, stimulating interaction of tumor cells with fibroblasts [12] and continues to systematically explore aspects of tumor-stroma interactions. Since some interactions require direct contact of the cells [12], a mixed co-culture, instead of transwell system in which cells contact only through the soluble signaling molecules, was used. A systematic overview of the heterotypic interactions effects of breast cancer and endothelial cells showed that interaction between breast cancer cells expressing CD44+/CD24- characteristics (e.g. Hs578T) and endothelial cells (HUVEC) induce genes with gene ontologies associated with increased proliferation and M-phase in the cell cycle. We demonstrated that genes induced due to the interaction of these cells include VEGF, FGF and other endothelial stimulatory factors. We managed to abrogate a large part of the stimulatory effect of Hs578T cells supernatant on HUVEC cells with bevacizumab. Since only a partial decrease of the stimulatory effect was observed, it may be suggested, that VEGF signaling is only one of many stimulatory loops between tumor and endothelial cells that might be targeted. It is probable, that genes present on our list are also responsible for drug resistance to anti-VEGF monotherapy as, presumably, they compensate the actions of blocked VEGF signaling. The signature was strongly correlated to other prognostic signatures, such as the “wound signature” and the “70-genes” signature and was able to stratify the breast cancer patients into two groups with different clinical outcomes. Since a predictive marker for anti-angiogenic therapies is missing, it can be speculated that the set of *in vitro*-designed

biomarkers, that already demonstrated prognostic power, is worth being checked in this context.

The development of drugs targeting the stroma or tumor-stroma interaction raises a need to develop specific markers that would allow assessing the efficacy of such approaches. Tailoring the therapy to only susceptible patients would allow increase the drug/cost effectiveness and reduce any eventual side effects. Testing stroma derived signatures, such as those presented in this thesis, could also be vital in predicting effectiveness and eventual side effects of therapies affecting the stroma or both the tumor and stroma.

Further perspectives

Advantages of primary cell culture

The direct tumor-stroma co-culture is a set-up mimicking the real tumor situation, which may be helpful in characterizing the signaling network between the cells. Based on our experience with the co-culture model that allowed direct cell-cell contact to mimic and simplify the situation of heterotypic interaction in a tumor, we think that this model would allow further exploration. The simple *in vitro* approach presented in this thesis allowed the description of stimulatory effects of tumor cells on endothelial cells and *vice versa*, analogically to mentioned co-cultivation of fibroblasts with tumor cells [12]. Since stromal cells in breast cancer differ from their normal analogues, the use of site-matching primary tumor and stromal cells would be an important addition to the use of the cell lines which might allow the observation of additional specific interaction effects. Specifically, co-culturing of primary carcinoma-associated endothelial cells with primary cancer cells, derived from the same patient, would be informative, since the cells would have the same genetic background. Primary human breast cancer cells would be the most reliable source of information, as they would carry all of the genetic (e.g. SNPs) and epigenetic (such as methylation) imprints that might be crucial for cancer progression. Furthermore, since it is speculated that the stroma is involved in tumorigenesis [28] exploration of the paracrine circuits in the cells from a single patient in whom cancer developed, might reveal more of the

specific interactions that allowed the formation of that cancer. Despite extensive attempts using several different culture conditions in parallel, unfortunately, the primary breast cancer cells could not be expanded extensively enough and kept proliferating to conduct such an experiment.

We speculated that, since gene expression changes accompanying tumor progression include the stromal compartment [28, 61], distinct differences between carcinoma-associated fibroblasts and normal fibroblasts only would allow building an effective prognosticator. In our laboratory we aimed to characterize the differences in gene expression changes between isolated primary breast fibroblasts and isolated carcinoma associated fibroblasts in paired specimens from the same patient and used the obtained gene expression signature as a prognosticator in the published datasets of early stage or advanced breast cancer. We managed to build up a collection of both tissues, however due to technical problems leading to subtle inconsistencies it was difficult to interpret the microarray gene expression profiling experiment and we were unable to formulate CAF-associated gene expression pattern which was significantly different from the primary normal breast fibroblasts associated gene expression pattern. We speculated that, since gene expression changes accompanying tumor progression include the stromal compartment [28, 61], distinct differences between carcinoma associated fibroblasts and normal fibroblasts would be enough to create a gene expression signature working as a valid prognostic marker in published gene expression datasets. The effectiveness of SDPP, published meanwhile by Finak *et al.* [120], which includes all types stromal cells, suggests that it is very likely as fibroblasts are significant fraction of tumor associated stroma and mediate most of paracrine signaling within tumor bulk [26]. Defining the hypothetical CAF specific signature and validating it as a predictor would allow estimating, if only all stroma features allow effective prognosis (like in SDPP) or if the signature coming from a single cell type (the hypothetical CAF signature) is enough to stratify patients with distinct clinical characteristics. Since carcinoma-associated fibroblasts retain their gene expression profiles when kept in cell culture [132], the use of single type of cells

(fibroblasts) in cell culture, in contrast to whole LCM dissected stroma, would be easier to monitor. It would also mean it would be easier to search for further aspects of stroma oriented therapies and stromal cells based prognosticators in breast cancer.

The collection of primary breast fibroblasts and carcinoma-associated fibroblasts in purified cell cultures gives us the opportunity to further characterize heterotypic interactions between these fibroblasts and other cells. Extending the previously published results [12] it would be possible to test if interactions of primary breast fibroblasts from healthy tissue with tumor cells differ from the interaction with CAFs and if the prognostic power of gene expression signatures derived from primary carcinoma associated cells would be increased in comparison to the one from cell lines.

Until now, all tested co-cultures and described interactions included tumor cells. It would also be interesting to check how the behavior of carcinoma-associated fibroblasts is modulated in the presence of other cells representing the tumor stroma. It is possible that the interaction between different tumor-associated stromal cells might affect cancer progression. Since all signaling loops existing between stromal cells affect the tumor associated stroma itself, the final resultant would affect the malignant epithelial cells. An easy to envision cascade of events is; secretion of growth factor in its non-mature form by one cell (e.g. endothelial) and cleavage of this factor to an active form by e.g. metaloproteinases secreted by CAFs, that are shown to express MMPs [47]. Active growth factors obtained this way would act on all cells present in the tumor milieu. For this purpose, it is possible to apply CAF cells and endothelial primary cells. Since it is already known that fibroblasts play an essential role in the angiogenic process through their production of extracellular matrix molecules and their release of essential growth factors [147], it would be interesting to test if tumor-associated fibroblasts act in a similar or even more potent manner on tumor associated endothelium. Unfortunately the volume of breast tissue that was obtained and the cell culture conditions, which were oriented towards fibroblast cell expansion, did not allow

the culture of primary carcinoma associated endothelial cells to verify this hypothesis.

Impact of hormones and growth factors on tumor-stroma interaction

Based on our experience with IGF-I we speculate that other growth factors have also a differential impact on tumor and tumor associated stromal cells. Their impact on carcinoma-associated fibroblasts could be systematically explored. An example of such an influence might be the recently demonstrated effect of Hedgehog (Hh) protein on cancer cells, as described by Yauch *et al.* [148]. Contradictory to the long accepted hypothesis, these authors showed that cancer cells *per se* do not respond to Hh and that Hh ligands fail to activate Hh signaling pathway in tumor epithelial cells. In contrast, they report the ligand–dependent activation of the Hh pathway in stromal cells. The Hh protein can affect the stroma by modulating the expression of different factors such as insulin-like growth factors and Wnt pathway components, which act on tumor cells and aid tumor progression. Some growth factors such as TGF-β, or hormones like estrogens, have been shown to have an indirect effect on tumor cells and act through fibroblasts. Since IGF-I stimulation causes fibroblasts to be in a more activated state, the effects of soluble factors secreted by primary fibroblasts in answer to IGF-I stimulation on tumor cells is definitely worth researching in the future. Defining similar loops in primary carcinoma-associated fibroblasts under the influence of other hormones and growth factors might aid the discovery of additional drug targets.

Combining the exploration of the tumor-stroma interaction induced gene expression changes with exploration of the hormone/growth factor impact on cells building tumor stroma is also a worthwhile aim. It has been shown that estrogens act directly and indirectly on tumor cells [14] and fibroblasts modulate the bioavailability of estrogens [67] by active induction of local estrogen synthesis. Furthermore, tumor necrosis factor *alpha* (TNFα) inhibits the differentiation of breast fibroblasts, and aromatase, the key enzyme in the biosynthesis of estrogen, is over-expressed in these undifferentiated fibroblasts,

producing large quantities of estrogen, which in turn influences the growth and progression of malignant epithelial cells [149]. Furthermore, estrogen modulates the synthesis of receptors for TNF in human adipose fibroblasts from breast tissue in a paracrine fashion, which may serve as a mechanism for the inhibition of adipocyte differentiation in breast cancer [149]. Further research should include whether the action of other factors is modulated in tumor-stroma co-culture. However, this would require the stimulation of tumor-stromal co-cultures cells with hormones/ growth factors in carefully chosen concentrations. Our preliminary data from the co-culture upon growth factor stimulation showed that the effects of stimuli are very potent compared to the effects of cell-cell interaction and can easily hide any effects of heterotypic interaction between the cells. In our first experiments, we treated the co-culture of equal numbers of human breast cancer cells MCF-7 with CCL-171 fibroblasts with 50 ng/ml of IGF-I. The gene expression profile of the co-culture stimulated with IGF-I was distinct from the not stimulated co-culture. Genes induced in the co-culture stimulated with IGF-I mostly represented genes highly up-regulated in one or the other cell line stimulated with IGF-I. This suggests that the effect we observed was rather an averaged value of expression in both cell types separately than the effect of IGF-I stimulation of the co-culture. Even though, some genes that were specifically induced in the co-culture upon IGF-I stimulation were observed. These included neurotensin and anterior gradient 2 homolog (Xenopus laevis), both known to be important in breast cancer [150-151]. The purpose of the investigation would be to define the global gene interactions between the cells in patients with increased signaling of the examined hormone/growth factor. Those hormones or growth factors might act at cancer cells in a paracrine manner through carcinoma-associated fibroblasts, which would presumably define new drug targets or allow a decrease in resistance to currently existing therapies. Radioprotection or resistance to chemotherapy may involve paracrine signaling via stroma cells, for example including aforementioned hedgehog stimulation of the stromal fibroblasts. Hh action on fibroblasts results in the expression of components of the IGF pathway by these cells [152]. IGF-I is known to provide

radioprotection and resistance to chemotherapeutic agents in tumor cells [153-154].

Differences in the stroma between different types of breast cancer

Since breast cancer is a very heterogeneous disease one could speculate that this heterogeneity is also apparent in the stroma. [88]. Is every molecular type of breast cancer connected to a specific type of the stroma that accompanies it? Since the molecular taxonomy of cancers was developed with gene expression profiles coming from the whole tumor bulk containing the stroma with multiple different cell types, such a scenario is very likely. Ropeman [137] demonstrated that a significant part of the gene expression profile comes from the stromal cells, it seems likely that part of the current taxonomy is based on the stroma derived gene expression patterns. Gene-expression pattern-based taxonomy of tumor stroma might form heterogeneous groups that match or do not match heterogeneous cancer groups [88] or for example the parallel estrogen receptor status of cancer cells that it had accompanied. If that was true, one should immediately check if this stroma derived taxonomy is correlated with patient performance in terms of disease progression and survival.

Furthermore, it would be possible to follow-up the patients who donated their tissue and compare their primary fibroblast gene expression profiles with clinical data. It would be of interest to see if unsupervised clustering of gene expression data of only carcinoma-associated fibroblasts was mirrored in clinical data ranging from receptor status of cancer cell up to patients survival.

Genetic and epigenetic changes of tumor associated stroma

We have shown the changes in gene expression on the mRNA level induced by tumor-stroma interaction and there is evidence that there are considerable differences between tumor-associated stroma and normal stroma of the breast. The mechanisms underlying these expression changes are not yet fully elucidated. Since there are tools to characterize small genetic changes, epigenetic modification and also micro RNAs which can significantly change the expression levels and bioavailability of factors influencing tumor and tumor-

stroma interaction, investigation of these modifications seems to be a worthwhile aim. The first step was already done by Allinen *et al.* who showed, with array comparative genomic hybridization (aCGH), that genetic changes, such as chromosomal gains or losses, are restrained only to epithelial cells [28]. The lack of severe genetic modifications does not exclude the presence of the other mild modifications like SNPs that can also influence expression levels and the bioactivity of different factors. Knowing whether other modifications are present in a defined fraction of stromal cells would allow testing of, for example, whether any specific SNPs influencing expression levels and functionality of proteins crucial for cancer cells's growth support. Since IGF-I, IGFBP1 and IGFBP3 gene polymorphisms have already been shown to be responsible for a change in circulating IGF-I levels [155], which is known to influence the course of breast cancer, one can speculate that other specific SNPs might also influence breast cancer progression. MicroRNAs (miRNAs) are a class of small noncoding RNAs that influence gene expression by targeting mRNAs and triggering either translation repression or RNA degradation. Aberrant expression of miRNAs contributes to carcinogenesis by promoting the expression of proto-oncogenes or by inhibiting the expression of tumor suppressor genes in different malignancies, including breast cancer [156]. Since miRNAs also play a role in the development of mouse mammary gland [156], it seems plausible that a changed profile of miRNA expression in the stroma would affect the behavior of the stromal cells and influence breast cancer progression. Presence of epigenetic changes in ductal carcinoma *in situ* and invasive breast carcinomas in the epithelial, in the myoepithelial cells and in the stromal fibroblasts in comparison to their analogues from normal breast tissue was already demonstrated by Hu *et al.* [157]. All of them, presumably, play a role in the maintenance of the abnormal cellular microenvironment in breast cancer [157]. It is likely that a more detailed study, involving more patients, would allow further characterization of epigenetic changes present in tumor-associated stroma. Defining the changes in stromal cells would allow, in future, analyzing how they are further modulated, or how they affect heterotypic interactions in the co-cultures of tumor and stromal cells.

In sillico derived prognostic markers

Since there are several stroma-derived gene expression patterns described it would be interesting the combine them and create a prognostic signature derived from stromal cells only. A similar analysis was already conducted for epithelial cells when Oncotype DX was created. I presume that combined signature created based on information extracted from the studies listed in the introduction would have a prognostic capacity for breast cancer patients as it covers gene expression alterations specific for all types of tumor associated stromal cells. Now much more detailed data (studies cited in the introduction) are available and the bioinformatics tools are more sophisticated. That could allow creating *in silico,* a better refined, only stroma-derived signature analogous to Oncotype DX, which could be an additional prognostic marker for breast cancer.

The recent trend toward improvement in breast cancer mortality rate is largely due to increased diagnosis of early stage disease, while therapeutic options for advanced stage carcinomas are still fairly limited [28]. Increased knowledge about the influence of stroma on tumor cells and paracrine regulatory circuits among various cell types in normal and cancerous breast tissue will allow the discovery of new drug therapeutic options. Our work on tumor-endothelium interactions reveals additional aspects of the cooperation of these cells. The description of effects of IGF-I on fibroblasts, comprising a significant part of the tumor associated stroma, show that external stimuli also have an impact on tumor stroma, which may be distinct from the impact on tumor cells. The effects of those factors on tumor stroma might influence signaling within the tumor milieu and modulate tumor progression. Moreover, as the stroma serves not only as a supportive environment for tumor progression but it can also provoke tumorgenicity in adjacent cells in the absence of a tumor, leading to the acquisition of genomic changes [9], it is worth considering these cells (stroma) as potientially important targets in cancer prevention by ablation of genes or signaling loops that might make breast tissue more prone to develop breast cancer. Additionally, predictive markers are needed to assess the efficacy of the already established anti-angiogenic therapies, as well as for the currently

developed anti-IGF-I therapies.

To summarize, our data highlights the importance of the tumor-endothelial cell interactions on global gene expression programs and shows that stromal cells also respond to external cues such as IGF-I. Genes induced in these systems have a prognostic capacity and might in addition be useful in deciding which is the most efficacious therapy for the individual concerned (personalized medicine).

Acknowledgements

I would like to express my gratitude to my supervisor, Dr. Martin Buess for his continuous, patient guidance, help and great degree of freedom that he offered me. This study could never have been achieved without him.

I would like to thank Professor Christoph Rochlitz for his advices during my time in his lab, valuable criticism to manuscripts, as well as for general help during my study.

I owe grateful thanks to the other members of the thesis evaluation committee, Prof. Nancy Hynes, Prof. Gerhard Christofori, and Professor Ulrich Certa who gave their valuable time to read, examine and evaluate this thesis.

I would like to express my thanks to Brigitte Vogel sharing her practical knowledge and experience throughout my studies. Thank you Brigitte for your great help and effective cooperation.

I would like to thank Florent Baty for sharing his practical knowledge and experience about R software.

I would like to thank all of the members of Friday Journal Club for their efforts to present papers, discussions and sharp comments that allowed me understanding that every “Nature” paper that glitters is not gold.

I express my sincere thanks to all my colleagues, Family and friends for their constant support and encouragement during my PhD studies.

Last but the most important, I would like to thank my dearest Magda for her constant support, help, motivation and being always the most important person for me.

Protocols

Protocol 1. Preparation of primary cell cultures from biopsies

1. Transport tumor tissue with some medium to the Petri dish. Aspirate the medium.
2. Separate tumor from surrounding tissue, skin and fat. Cut the tumor in as small as possible parts.
3. Transfer fragmentized tumor tissue into a sterile bottle and treat with collagenase / DRNase – mix for 1 hour (37°C).
4. Place the material on a sieve and mince the tumor material with light press of piston. Wash the sieve with PBS.
5. Place the material in a tube and centrifuge 1300 r.p.m. / 10 min.
6. Aspirate supernatant and wash the pellet with PBS. Centrifuge again and resuspend in PBS.
7. Count the cells and set a cell culture in an appropriate dish.
8. In parallel, proceed the normal tissue according to the same procedure.

Protocol 2. Coating of the plates with Poly – L – Lysine

1. Prepare stock solution of Poly-L-lysine hydrobromide (0.5 mg/mL) in water.
2. Dilute the stock solution 1:30 (v/v) in sterile water.
3. Cover the surfaces that are to be coated with the Poly-L-lysine solution and incubate for one hour (37ºC).
4. Wash the surface twice with ice – cold sterile water and let it dry in a sterile environment.

Protocol 3. Cell culture media preparation

Cell culture media were prepared by admixing RPMI 1640 15% FCS 1:1 (v/v.) with F12 nutrient mix.

RPMI 1640 15% FBS (v/v).		**F12 nutrient 10% FBS (v/v.)**	
RPMI 1640 medium	- 400mL.	F 12 medium	- 450mL.
FBS (heat inactivated)	- 75mL.	FBS (heat inactivated)	- 50mL.
Penicillin/Streptomycin	- 5mL.	Penicillin/Streptomycin	- 5mL.
Puryvat	- 5mL.	Puryvat	- 5mL.
L-glutamin	-10mL.		
Non essential aminoacids	- 5mL.		
Vitamines mix	- 5mL.		
2-mercapto-ethanol	- 0.5mL.		

Protocol 4. Microbeads separation protocols

The procedure was conducted according to users' manual protocol (Miltenyi Biotec; Anti-Fibroblast MicroBeads - Part Number 130-050-601).

Protocol 5. WST -1 staining

1. Set up the cell culture, that is to be measured, in a 96 well plate.
2. At the day of measurement prepare 10% v/v solution of WST-1 in cell culture medium without additives (no FCS, no goodies, with antibiotics).
3. Aspirate the cell culture medium from the wells of the 96 well plate with an automatic pipette.
4. Wash the wells ones with PBS. Work gently not to scratch the cell monolayer.
5. In the dark, add 70 µL of 10% WST-1 solution to every well of the 96 well plate.
6. Incubate for 1h at 37ºC in a cell culture incubator.
7. Measure the absorbance of formazan dye produced by metabolically active cells at wavelength 450 nm.

Protocol 6. Amplification of RNA

The procedure was conducted according to users' manual protocol (MessageAmp™ Kit - Part Number AM1750) for both; sample RNA and reference RNA.

Protocol 7. Microarrays

Post processing of Oligo Arrays

1. UV cross-link printed DNA onto glass substrate.
2. Prehybridize microarrays with cross-linked DNA in 5 x SSC, 0.1 mg/mL BSA, 0.1 % SDS solution 1h at 42ºC.
3. Rinse the microarrays twice in 0.1 x SSC solution for 5 minutes in a room temperature (RT).
4. Rinse the microarrays in deionized water for 30 seconds at RT.
5. Spin-dry the microarrays and them the same day.

The aRNA/doping controls mix preparation

1. Prepare 8µg sample amplified RNA (aRNA) and reference aRNA.
2. Mix 8µg of sample aRNA with 5µl of Cy5 stained doping controls mix.
3. Mix 8µg of reference aRNA with 5µl of Cy3 stained doping controls mix.
4. Vacuum dry the samples and the references.
5. Add 9µl of a coupling buffer to the dried aRNA/doping controls mix and reference aRNA/doping control mix.

Staining the sample aRNA and the reference aRNA

1. Prepare 11µl of Cy5 dye solution in DMSO for aRNA/doping controls mix.
2. Prepare 11µl of Cy3 dye solution in DMSO for reference aRNA/doping controls mix
3. Store in dark for maximum one hour.
1. Add 11µl of prepared Cy5 dye to aRNA/doping controls mix.
2. Add 11µl of prepared Cy3 dye to reference aRNA/doping controls mix.
3. Incubate 30 minutes at RT in the dark.
4. Quench the reaction by adding 4.5µl Hydroxylamine to aRNA/doping controls mix and reference aRNA/doping controls mix (15 minutes at RT in the dark).

aRNA purification

1. Add 105µl of aRNA Binding Buffer to aRNA/doping controls mix and reference aRNA/doping controls mix.
2. Add 75µl of ACS grade 100% ethanol to aRNA/doping controls mix and reference aRNA/doping controls mix.

3. Pool sample aRNA/doping controls mix and reference aRNA/doping controls mix in a “Labeled aRNA filter cartridge”. Centrifuge for 1 minute at 10,000x g.
4. Wash with 500µl of aRNA Wash Buffer.
5. Centrifuge for 1 minute at 10,000x g.
6. Centrifuge for an additional 1 min at 10,000x g to remove trace amounts of ethanol.
7. Transfer cartridge to a fresh “Labeled aRNA Elution Tube”.
8. Incubate at RT for 2 minutes with 15µl of preheated water (50-60^{o}C).
9. Centrifuge for 1.5 min at 10,000x g and repeat the elution.

Fragmentation of probe prior to hybridization

1. Bring the labeled aRNA probe to 27 µl with Nuclease-free water.
2. Fragmentize the probe with 3µl of 10xFragmentation buffer at 70ºC for 15 minutes.
3. Block the reaction with 3µl of Stop solution.
4. Bring the probe to the volume of 38µl with 5µl of water.

Probe preparation

1. Mix the probe with 6 µl of PolyA/Cot1/tRNA mix (10 ug/µl each).
2. Mix the probe with 9.35 µl of 20 x SSC solution.
3. Add 1 µl of HEPES buffer (1M).
4. Add 1.65 µl of 10% SDS.
5. Denature the probe for 5 minutes at 70ºC and centrifuge at 10,000x g for 5 minutes.
6. Put warm probe on microarray slide and close them in a hybridization chamber.
7. Incubate in the chamber at 65ºC for 16 hours.

Microarrays washing

1. Take out the hybridization chamber from the water bath.
2. Dry the chamber exterior, unscrew chamber and take out the microarray and gently allow cover slip to fall off in 2xSSC, 0.1% SDS solution.
3. Transfer array to a slide rack and wash in a fresh 2xSSC, 0.1% SDS solution.
4. Transfer slide rack to 60ºC, 2xSSC, 0.1% SDS solution and agitate for 1.5 minutes.
5. Dip microarray in 2xSSC solution and swirl around to get rid of SDS
6. Transfer microarrays to slide rack in the 1xSSC solution and agitate for 1.5 minutes.
7. Transfer the entire slide rack to the 0.2xSSC solution and agitate for 1.5 minutes.

8. Spin dry microarray at 600 RPM for 5 min.
9. Scan the microarray immediately.

Protocol 7. Microarray data retrieval

1. The row data were retrieved and averaged by : BIOSEQUENCE_ID for Log(base2) of Cy5/Cy3 normalized ratio (Mean).
2. Data were considered valid when spot was not flagged by experimenter as an artifact and regression correlation was bigger than 0.6 and channel 1 normalized ratio (mean intensity over median background intensity) was bigger than 1.5 or channel 2 normalized ratio (mean intensity over median background intensity) was bigger than 1.5.
3. Genes and arrays were centered by mean and in case of stimulation experiments gene wise centered by not stimulated sample.
4. Only genes with more than 80% good data present were analyzed.

Protocol 8. Chosen R-scripts

Function – Extract selected genes from a dataset.

This script was used to extract the gene expression values for the signature (prepared as a list of unique gene IDs) from the dataset containing gene expression data of the patients (like NKI dataset). Genes present as duplicates in the dataset were as well extracted in duplicates.

```
match2 <- function(g1, x){
  st1 <- strsplit(g1, "\|", fixed=T)
  l1 <- lapply(st1, function(z) lapply(as.list(z), function(zz)
if(zz!=""){grep(paste("|",zz,"|", sep=""), rownames(x), extended=FALSE)}))
  x[unique(unlist(l1)),]
}
```

Function – Average extracted genes.

This script was used to average the gene expression data for the genes that had duplicates within dataset. The effect was that every gene ID present in the signature had only one (averaged) value for an each patient.

```
averageGenes <- function(x){
      rn <- rownames(x)
      rn2 <- strsplit(rn, "\|", fixed=T)
      facrn3 <- as.factor(unlist(lapply(rn2, function(z) ifelse(z[1]=="",
ifelse(z[2]=="", return(z[3]), return(z[2])), return(z[1])))))
      spx <- split(x, facrn3)
      x2 <- t(sapply(spx, colMeans, na.rm = TRUE))
      return(x2)
}
```

Function - Survival calculation.

This script was used to conduct the survival analysis. It uses the package "survival" being a part of R library: "survival".

```
survival <- function(x){
library(survival)
x1 <- survfit(Surv(SurvTime, Status)~Category, data=x)
x2 <- coxph(Surv(SurvTime, Status)~Category, data=x)
x3 <- survdiff(Surv(SurvTime, Status)~Category, data=x) plot(x1, col=1:2, lty=1,
xlab="Time [Years]",ylab="Probability")
legend("bottom",title = "Legend:", legend=c("left", "right"), col=1:2, lty=1, bty="n")
print(summary(x2)$logtest["pvalue"])
print(x3)
}
```

Multivariable analysis of survival.

This script was used to conduct the multivariable survival analysis. It uses the package "survival" being a part of R library: "survival".

```
multivariable <- function(x){
library(survival)
fit1 <- survfit(Surv(SurvTime, Status) ~ Category + ER, data=x)
plot(fit1, col=1:nlevels(x$Category:x$ER), lty=1, lwd = 3, mark.time=F)
legend("bottomleft", legend=levels(x$Category:x$ER), lty=1, lwd=3,
col=1:nlevels(x$Category:x$ER))

stat1 <- coxph(Surv(SurvTime, Status) ~ Category + ER, data=x)

print(summary(stat1)$logtest["pvalue"])
}
```

Abbreviations

aCGH - array comparative genomic hybridization.

BRCA1 - breast cancer 1, early onset

BRCA2 - breast cancer type 2 susceptibility protein.

CAF – carcinoma associated fibroblasts.

CCL5 (RANTES) - chemokine (C-C motif) ligand 5.

CD – cluster of differentiation.

CGH - comparative genomic hybridization.

CSC – cancer stem cell.

CXCL12 (SDF1) - stromall cell-derived factor 1.

CXCR4 - C-X-C chemokine receptor type 4.

DCIS - ductal carcinoma in situ.

DMFS - distant metastasis free survival.

DNA - Deoxyribonucleic acid

DTF - desmoid-type fibromatosis

ECM - extracellular matrix.

EMT - epithelial to mesenchymal transition.

ER – estrogen receptor.

ErbB2 (HER2/neu, ErbB-2, ERBB2) Human Epidermal growth factor Receptor 2.

ESA - epithelial specific antigen.

FAP - Fibroblast activation protein alpha.

FGF - Fibroblast growth factor.

FSP1 (mts1) - fibroblast secreted protein-1.

HEEBO - Human Exonic Evidence Based Oligonucleotide microarrays.

HEYL - hairy/enhancer-of-split.

HGF - hepatocyte growth factor.

Hh – Hedgehog.

hMSCs - human mesenchymal stem cells.

HNSCCS - head and neck squamous cell carcinoma.

HUVEC - Human Umbilical Vein Endothelial Cells.

IDC – invasive ductal carcinoma.

IGFBP1 - insulin like growth factor binding protein one.

IGFBP3 - insulin like growth factor binding protein three.

IGF-I – insulin like growth factor one.

IGFIR – insulin like growth factor receptor one.

IL-8 – interleukin 8.

IR – insulin receptor.

IRG - interferon response genes.

LCM - laser capture microdissected.

LOXL2 - lysyl oxidase-like 2.

MAPK - mitogen-activated protein (MAP) kinases.

MINDACT - Microarray In Node-negative and 1 to 3 positive lymph node Disease may Avoid ChemoTherapy.

MMP – metaloproteinases.

MMP-1 – metalloproteinase 1.

MMP-13 – metalloproteinase 13.

MMP-2 – metalloproteinase 2.

p53 - tumor protein 53.

PI3K - Phosphoinositide 3-kinases

POSTN – periostin.

PRL – prolactin.

PRL3 - Protein-tyrosine phosphatase type IVa, member 3.

RNA - Ribonucleic acid.

SAGE - serial analysis of gene expression.

Sca1 - stem cell antigen one.

SDPP - stroma derived prognostic predictor.

SNAIL1 - Snail 1 Drosophila homolog.

SNP - single nucleotide polymorphism.

SPARC - secreted protein, acidic, cysteine-rich.

STAT1 - signal transducer and activator of transcription 1.

STF - solitary fibrous tumors

TGF-β - transforming growth factor beta.

TLL2 - tolloid-like 2.

VEGF - vascular endothelial growth factor.

VEGF-A - vascular endothelial growth factor-A.

αSMA - alpha smooth muscle actin.

References

1. World Health Organization: **http://www.who.int/cancer/en/**.
2. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ: **Cancer statistics, 2009**. *CA Cancer J Clin* 2009, **59**(4):225-249.
3. Alpers CE, Wellings SR: **The prevalence of carcinoma in situ in normal and cancer-associated breasts**. *Hum Pathol* 1985, **16**(8):796-807.
4. Nielsen M, Thomsen JL, Primdahl S, Dyreborg U, Andersen JA: **Breast cancer and atypia among young and middle-aged women: a study of 110 medicolegal autopsies**. *Br J Cancer* 1987, **56**(6):814-819.
5. Thiery JP: **Epithelial-mesenchymal transitions in development and pathologies**. *Current Opinion in Cell Biology* 2003, **15**(6):740-746.
6. Folkman J: **What is the evidence that tumors are angiogenesis dependent?** *J Natl Cancer Inst* 1990, **82**(1):4-6.
7. Brabletz T, Jung A, Reu S, Porzner M, Hlubek F, Kunz-Schughart LA, Knuechel R, Kirchner T: **Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment**. *Proc Natl Acad Sci U S A* 2001, **98**(18):10356-10361.
8. Al-Hajj M, Clarke MF: **Self-renewal and solid tumor stem cells**. *Oncogene* 2004, **23**(43):7274-7282.
9. Kim JB, Stein R, O'Hare MJ: **Tumour-stromal interactions in breast cancer: the role of stroma in tumourigenesis**. *Tumour Biol* 2005, **26**(4):173-185.
10. Karnoub AE, Dash AB, Vo AP, Sullivan A, Brooks MW, Bell GW, Richardson AL, Polyak K, Tubo R, Weinberg RA: **Mesenchymal stem cells within tumour stroma promote breast cancer metastasis**. *Nature* 2007, **449**(7162):557-563.
11. Orimo A, Gupta PB, Sgroi DC, Arenzana-Seisdedos F, Delaunay T, Naeem R, Carey VJ, Richardson AL, Weinberg RA: **Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion**. *Cell* 2005, **121**(3):335-348.
12. Buess M, Nuyten DS, Hastie T, Nielsen T, Pesich R, Brown PO: **Characterization of heterotypic interaction effects in vitro to deconvolute global gene expression profiles in cancer**. *Genome Biol* 2007, **8**(9):R191.
13. Bierie B, Moses HL: **TGF-beta and cancer**. *Cytokine Growth Factor Rev* 2006, **17**(1-2):29-40.
14. Clemons M, Goss P: **Estrogen and the risk of breast cancer**. *N Engl J Med* 2001, **344**(4):276-285.
15. Ingram DM, Nottage EM, Roberts AN: **Prolactin and breast cancer risk**. *Med J Aust* 1990, **153**(8):469-473.
16. Tworoger SS, Hankinson SE: **Prolactin and breast cancer risk**. *Cancer Lett* 2006, **243**(2):160-169.
17. Reya T, Morrison SJ, Clarke MF, Weissman IL: **Stem cells, cancer, and cancer stem cells**. *Nature* 2001, **414**(6859):105-111.

18. Cicalese A, Bonizzi G, Pasi CE, Faretta M, Ronzoni S, Giulini B, Brisken C, Minucci S, Di Fiore PP, Pelicci PG: **The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells**. *Cell* 2009, **138**(6):1083-1095.
19. Bonnet D, Dick JE: **Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell**. *Nat Med* 1997, **3**(7):730-737.
20. Ponti D, Costa A, Zaffaroni N, Pratesi G, Petrangolini G, Coradini D, Pilotti S, Pierotti MA, Daidone MG: **Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties**. *Cancer Res* 2005, **65**(13):5506-5511.
21. Shipitsin M, Campbell LL, Argani P, Weremowicz S, Bloushtain-Qimron N, Yao J, Nikolskaya T, Serebryiskaya T, Beroukhim R, Hu M *et al*: **Molecular definition of breast tumor heterogeneity**. *Cancer Cell* 2007, **11**(3):259-273.
22. Charafe-Jauffret E, Ginestier C, Iovino F, Wicinski J, Cervera N, Finetti P, Hur MH, Diebel ME, Monville F, Dutcher J *et al*: **Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature**. *Cancer Res* 2009, **69**(4):1302-1313.
23. Rich JN, Bao S: **Chemotherapy and cancer stem cells**. *Cell Stem Cell* 2007, **1**(4):353-355.
24. Todaro M, Alea MP, Di Stefano AB, Cammareri P, Vermeulen L, Iovino F, Tripodo C, Russo A, Gulotta G, Medema JP *et al*: **Colon cancer stem cells dictate tumor growth and resist cell death by production of interleukin-4**. *Cell Stem Cell* 2007, **1**(4):389-402.
25. Klonisch T, Wiechec E, Hombach-Klonisch S, Ande SR, Wesselborg S, Schulze-Osthoff K, Los M: **Cancer stem cell markers in common cancers - therapeutic implications**. *Trends Mol Med* 2008, **14**(10):450-460.
26. Elenbaas B, Weinberg RA: **Heterotypic signaling between epithelial tumor cells and fibroblasts in carcinoma formation**. *Exp Cell Res* 2001, **264**(1):169-184.
27. Brisken C, Duss S: **Stem cells and the stem cell niche in the breast: an integrated hormonal and developmental perspective**. *Stem Cell Rev* 2007, **3**(2):147-156.
28. Allinen M, Beroukhim R, Cai L, Brennan C, Lahti-Domenici J, Huang H, Porter D, Hu M, Chin L, Richardson A *et al*: **Molecular characterization of the tumor microenvironment in breast cancer**. *Cancer Cell* 2004, **6**(1):17-32.
29. Egeblad M, Littlepage LE, Werb Z: **The fibroblastic coconspirator in cancer progression**. *Cold Spring Harb Symp Quant Biol* 2005, **70**:383-388.
30. Iiizumi M, Mohinta S, Bandyopadhyay S, Watabe K: **Tumor-endothelial cell interactions: therapeutic potential**. *Microvasc Res* 2007, **74**(2-3):114-120.

31. Honn KV, Tang DG: **Adhesion molecules and tumor cell interaction with endothelium and subendothelial matrix**. *Cancer Metastasis Rev* 1992, **11**(3-4):353-375.
32. Zumsteg A, Christofori G: **Corrupt policemen: inflammatory cells promote tumor angiogenesis**. *Curr Opin Oncol* 2009, **21**(1):60-70.
33. Coussens LM, Werb Z: **Inflammation and cancer**. *Nature* 2002, **420**(6917):860-867.
34. Barcellos-Hoff MH, Ravani SA: **Irradiated mammary gland stroma promotes the expression of tumorigenic potential by unirradiated epithelial cells**. *Cancer Res* 2000, **60**(5):1254-1260.
35. Bissell MJ, Weaver VM, Lelievre SA, Wang F, Petersen OW, Schmeichel KL: **Tissue structure, nuclear organization, and gene expression in normal and malignant breast**. *Cancer Res* 1999, **59**(7 Suppl):1757-1763s; discussion 1763s-1764s.
36. Deng G, Lu Y, Zlotnikov G, Thor AD, Smith HS: **Loss of heterozygosity in normal tissue adjacent to breast carcinomas**. *Science* 1996, **274**(5295):2057-2059.
37. Hu M, Yao J, Carroll DK, Weremowicz S, Chen H, Carrasco D, Richardson A, Violette S, Nikolskaya T, Nikolsky Y *et al*: **Regulation of In Situ to Invasive Breast Carcinoma Transition**. 2008, **13**(5):394-406.
38. Bhowmick NA, Chytil A, Plieth D, Gorska AE, Dumont N, Shappell S, Washington MK, Neilson EG, Moses HL: **TGF-beta signaling in fibroblasts modulates the oncogenic potential of adjacent epithelia**. *Science* 2004, **303**(5659):848-851.
39. Kuperwasser C, Chavarria T, Wu M, Magrane G, Gray JW, Carey L, Richardson A, Weinberg RA: **Reconstruction of functionally normal and malignant human breast tissues in mice**. *Proc Natl Acad Sci U S A* 2004, **101**(14):4966-4971.
40. Dong-Le Bourhis X, Berthois Y, Millot G, Degeorges A, Sylvi M, Martin PM, Calvo F: **Effect of stromal and epithelial cells derived from normal and tumorous breast tissue on the proliferation of human breast cancer cell lines in co-culture**. *Int J Cancer* 1997, **71**(1):42-48.
41. van Roozendaal KE, Klijn JG, van Ooijen B, Claassen C, Eggermont AM, Henzen-Logmans SC, Foekens JA: **Differential regulation of breast tumor cell proliferation by stromal fibroblasts of various breast tissue sources**. *Int J Cancer* 1996, **65**(1):120-125.
42. McAllister SS, Gifford AM, Greiner AL, Kelleher SP, Saelzler MP, Ince TA, Reinhardt F, Harris LN, Hylander BL, Repasky EA *et al*: **Systemic endocrine instigation of indolent tumor growth requires osteopontin**. *Cell* 2008, **133**(6):994-1005.
43. Shekhar MP, Werdell J, Santner SJ, Pauley RJ, Tait L: **Breast stroma plays a dominant regulatory role in breast epithelial growth and differentiation: implications for tumor development and progression**. *Cancer Res* 2001, **61**(4):1320-1326.
44. Grum-Schwensen B, Klingelhofer J, Berg CH, El-Naaman C, Grigorian M, Lukanidin E, Ambartsumian N: **Suppression of tumor development and**

metastasis formation in mice lacking the S100A4(mts1) gene. *Cancer Res* 2005, **65**(9):3772-3780.
45. Bleul CC, Fuhlbrigge RC, Casasnovas JM, Aiuti A, Springer TA: **A highly efficacious lymphocyte chemoattractant, stromal cell-derived factor 1 (SDF-1)**. *J Exp Med* 1996, **184**(3):1101-1109.
46. Vacek PM, Geller BM: **A prospective study of breast cancer risk using routine mammographic breast density measurements**. *Cancer Epidemiol Biomarkers Prev* 2004, **13**(5):715-722.
47. Nielsen BS, Rank F, Lopez JM, Balbin M, Vizoso F, Lund LR, Dano K, Lopez-Otin C: **Collagenase-3 expression in breast myofibroblasts as a molecular marker of transition of ductal carcinoma in situ lesions to invasive ductal carcinomas**. *Cancer Res* 2001, **61**(19):7091-7100.
48. Uria JA, Stahle-Backdahl M, Seiki M, Fueyo A, Lopez-Otin C: **Regulation of collagenase-3 expression in human breast carcinomas is mediated by stromal-epithelial cell interactions**. *Cancer Res* 1997, **57**(21):4882-4888.
49. Balduyck M, Zerimech F, Gouyer V, Lemaire R, Hemon B, Grard G, Thiebaut C, Lemaire V, Dacquembronne E, Duhem T *et al*: **Specific expression of matrix metalloproteinases 1, 3, 9 and 13 associated with invasiveness of breast cancer cells in vitro**. *Clin Exp Metastasis* 2000, **18**(2):171-178.
50. Zhang B, Cao X, Liu Y, Cao W, Zhang F, Zhang S, Li H, Ning L, Fu L, Niu Y *et al*: **Tumor-derived matrix metalloproteinase-13 (MMP-13) correlates with poor prognoses of invasive breast cancer**. *BMC Cancer* 2008, **8**:83.
51. Holmgren L, O'Reilly MS, Folkman J: **Dormancy of micrometastases: balanced proliferation and apoptosis in the presence of angiogenesis suppression**. *Nat Med* 1995, **1**(2):149-153.
52. Ferrara N: **VEGF as a therapeutic target in cancer**. *Oncology* 2005, **69 Suppl 3**:11-16.
53. Parker BS, Argani P, Cook BP, Liangfeng H, Chartrand SD, Zhang M, Saha S, Bardelli A, Jiang Y, St Martin TB *et al*: **Alterations in vascular gene expression in invasive breast carcinoma**. *Cancer Res* 2004, **64**(21):7857-7866.
54. Okamoto H, Ohigashi H, Nakamori S, Ishikawa O, Imaoka S, Mukai M, Kusama T, Fujii H, Matsumoto Y, Akedo H: **Reciprocal functions of liver tumor cells and endothelial cells. Involvement of endothelial cell migration and tumor cell proliferation at a primary site in distant metastasis**. *Eur Surg Res* 2000, **32**(6):374-379.
55. Li L, Neaves WB: **Normal stem cells and cancer stem cells: the niche matters**. *Cancer Res* 2006, **66**(9):4553-4557.
56. Paez-Ribes M, Allen E, Hudock J, Takeda T, Okuyama H, Vinals F, Inoue M, Bergers G, Hanahan D, Casanovas O: **Antiangiogenic therapy elicits malignant progression of tumors to increased local invasion and distant metastasis**. *Cancer Cell* 2009, **15**(3):220-231.

57. Crawford Y, Kasman I, Yu L, Zhong C, Wu X, Modrusan Z, Kaminker J, Ferrara N: **PDGF-C mediates the angiogenic and tumorigenic properties of fibroblasts associated with tumors refractory to anti-VEGF treatment**. *Cancer Cell* 2009, **15**(1):21-34.
58. Rhodes LV, Muir SE, Elliott S, Guillot LM, Antoon JW, Penfornis P, Tilghman SL, Salvo VA, Fonseca JP, Lacey MR *et al*: **Adult human mesenchymal stem cells enhance breast tumorigenesis and promote hormone independence**. *Breast Cancer Res Treat* 2009.
59. Rajagopalan H, Nowak MA, Vogelstein B, Lengauer C: **The significance of unstable chromosomes in colorectal cancer**. *Nat Rev Cancer* 2003, **3**(9):695-701.
60. Weinberg RA: **Coevolution in the tumor microenvironment**. *Nat Genet* 2008, **40**(5):494-495.
61. Ma XJ, Dahiya S, Richardson E, Erlander M, Sgroi DC: **Gene expression profiling of the tumor microenvironment during breast cancer progression**. *Breast Cancer Res* 2009, **11**(1):R7.
62. Hankinson SE, Willett WC, Michaud DS, Manson JE, Colditz GA, Longcope C, Rosner B, Speizer FE: **Plasma prolactin levels and subsequent risk of breast cancer in postmenopausal women**. *J Natl Cancer Inst* 1999, **91**(7):629-634.
63. Toniolo PG: **Endogenous estrogens and breast cancer risk: the case for prospective cohort studies**. *Environ Health Perspect* 1997, **105 Suppl 3**:587-592.
64. Beatson GT: **On the treatment of inoperable cases of carcinoma of the mamma: suggestions for a new method of treatment, with illustrative cases.** *Lancet* 1896, **2**:104-107.
65. **Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Collaborative Group on Hormonal Factors in Breast Cancer**. *Lancet* 1997, **350**(9084):1047-1059.
66. Bentzon N, During M, Rasmussen BB, Mouridsen H, Kroman N: **Prognostic effect of estrogen receptor status across age in primary breast cancer**. *Int J Cancer* 2008, **122**(5):1089-1094.
67. Yamaguchi Y, Hayashi S: **Estrogen-related cancer microenvironment of breast carcinoma**. *Endocr J* 2009, **56**(1):1-7.
68. Malarkey WB, Kennedy M, Allred LE, Milo G: **Physiological concentrations of prolactin can promote the growth of human breast tumor cells in culture**. *J Clin Endocrinol Metab* 1983, **56**(4):673-677.
69. Peyrat JP, Djiane J, Bonneterre J, Vandewalle B, Vennin P, Delobelle A, Depadt G, Lefebvre J: **Stimulation of DNA synthesis by prolactin in human breast tumor explants. Relation to prolactin receptors**. *Anticancer Res* 1984, **4**(4-5):257-261.
70. Simon WE, Albrecht M, Trams G, Dietel M, Holzel F: **In vitro growth promotion of human mammary carcinoma cells by steroid hormones, tamoxifen, and prolactin**. *J Natl Cancer Inst* 1984, **73**(2):313-321.

71. Welsch CW, Nagasawa H: **Prolactin and murine mammary tumorigenesis: a review**. *Cancer Res* 1977, **37**(4):951-963.
72. Tworoger SS, Sluss P, Hankinson SE: **Association between plasma prolactin concentrations and risk of breast cancer among predominately premenopausal women**. *Cancer Res* 2006, **66**(4):2476-2482.
73. Dupont J, Khan J, Qu BH, Metzler P, Helman L, LeRoith D: **Insulin and IGF-1 induce different patterns of gene expression in mouse fibroblast NIH-3T3 cells: identification by cDNA microarray analysis**. *Endocrinology* 2001, **142**(11):4969-4975.
74. Nandeesha H: **Insulin: a novel agent in the pathogenesis of prostate cancer**. *Int Urol Nephrol* 2009, **41**(2):267-272.
75. Igarashi M, Yamaguchi H, Hirata A, Daimon M, Tominaga M, Kato T: **Insulin activates p38 mitogen-activated protein (MAP) kinase via a MAP kinase kinase (MKK) 3/MKK 6 pathway in vascular smooth muscle cells**. *Eur J Clin Invest* 2000, **30**(8):668-677.
76. Lann D, LeRoith D: **The role of endocrine insulin-like growth factor-I and insulin in breast cancer**. *J Mammary Gland Biol Neoplasia* 2008, **13**(4):371-379.
77. Panno ML, Salerno M, Pezzi V, Sisci D, Maggiolini M, Mauro L, Morrone EG, Ando S: **Effect of oestradiol and insulin on the proliferative pattern and on oestrogen and progesterone receptor contents in MCF-7 cells**. *J Cancer Res Clin Oncol* 1996, **122**(12):745-749.
78. McTernan PG, Anwar A, Eggo MC, Barnett AH, Stewart PM, Kumar S: **Gender differences in the regulation of P450 aromatase expression and activity in human adipose tissue**. *Int J Obes Relat Metab Disord* 2000, **24**(7):875-881.
79. Karp DD, Paz-Ares LG, Novello S, Haluska P, Garland L, Cardenal F, Blakely LJ, Eisenberg PD, Langer CJ, Blumenschein G, Jr. *et al*: **Phase II study of the anti-insulin-like growth factor type 1 receptor antibody CP-751,871 in combination with paclitaxel and carboplatin in previously untreated, locally advanced, or metastatic non-small-cell lung cancer**. *J Clin Oncol* 2009, **27**(15):2516-2522.
80. Scarpa RC, Carraway RE, Cochrane DE: **Insulin-like growth factor (IGF) induced proliferation of human lung fibroblasts is enhanced by neurotensin**. *Peptides* 2005, **26**(11):2201-2210.
81. Yateman ME, Claffey DC, Cwyfan Hughes SC, Frost VJ, Wass JA, Holly JM: **Cytokines modulate the sensitivity of human fibroblasts to stimulation with insulin-like growth factor-I (IGF-I) by altering endogenous IGF-binding protein production**. *J Endocrinol* 1993, **137**(1):151-159.
82. Johansson GS, Chisalita SI, Arnqvist HJ: **Human microvascular endothelial cells are sensitive to IGF-I but resistant to insulin at the receptor level**. *Mol Cell Endocrinol* 2008, **296**(1-2):58-63.

83. Kulesh DA, Clive DR, Zarlenga DS, Greene JJ: **Identification of interferon-modulated proliferation-related cDNA sequences**. *Proc Natl Acad Sci U S A* 1987, **84**(23):8453-8457.
84. Schena M, Shalon D, Davis RW, Brown PO: **Quantitative monitoring of gene expression patterns with a complementary DNA microarray**. *Science* 1995, **270**(5235):467-470.
85. [http://www.nimblegen.com/products/exp/index.html]
86. http://www.microarray.org/. In.; 2009.
87. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA *et al*: **Molecular portraits of human breast tumours**. *Nature* 2000, **406**(6797):747-752.
88. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS *et al*: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications**. *Proc Natl Acad Sci U S A* 2001, **98**(19):10869-10874.
89. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S *et al*: **Repeated observation of breast tumor subtypes in independent gene expression data sets**. *Proc Natl Acad Sci U S A* 2003, **100**(14):8418-8423.
90. Garber ME, Troyanskaya OG, Schluens K, Petersen S, Thaesler Z, Pacyna-Gengelbach M, van de Rijn M, Rosen GD, Perou CM, Whyte RI *et al*: **Diversity of gene expression in adenocarcinoma of the lung**. *Proc Natl Acad Sci U S A* 2001, **98**(24):13784-13789.
91. Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, Vasa P, Ladd C, Beheshti J, Bueno R, Gillette M *et al*: **Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses**. *Proc Natl Acad Sci U S A* 2001, **98**(24):13790-13795.
92. Nielsen TO, West RB, Linn SC, Alter O, Knowling MA, O'Connell JX, Zhu S, Fero M, Sherlock G, Pollack JR *et al*: **Molecular characterisation of soft tissue tumours: a gene expression study**. *Lancet* 2002, **359**(9314):1301-1307.
93. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X *et al*: **Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling**. *Nature* 2000, **403**(6769):503-511.
94. Chen X, Cheung ST, So S, Fan ST, Barry C, Higgins J, Lai KM, Ji J, Dudoit S, Ng IO *et al*: **Gene expression patterns in human liver cancers**. *Mol Biol Cell* 2002, **13**(6):1929-1939.
95. Gruvberger S, Ringner M, Chen Y, Panavally S, Saal LH, Borg A, Ferno M, Peterson C, Meltzer PS: **Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns**. *Cancer Res* 2001, **61**(16):5979-5984.
96. Martin KJ, Kritzman BM, Price LM, Koh B, Kwan CP, Zhang X, Mackay A, O'Hare MJ, Kaelin CM, Mutter GL *et al*: **Linking gene expression**

patterns to therapeutic groups in breast cancer. *Cancer Res* 2000, **60**(8):2232-2238.
97. Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JC *et al*: **Distinctive gene expression patterns in human mammary epithelial cells and breast cancers**. *Proc Natl Acad Sci U S A* 1999, **96**(16):9212-9217.
98. West M, Blanchette C, Dressman H, Huang E, Ishida S, Spang R, Zuzan H, Olson JA, Jr., Marks JR, Nevins JR: **Predicting the clinical status of human breast cancer by using gene expression profiles**. *Proc Natl Acad Sci U S A* 2001, **98**(20):11462-11467.
99. Zajchowski DA, Bartholdi MF, Gong Y, Webster L, Liu HL, Munishkin A, Beauheim C, Harvey S, Ethier SP, Johnson PH: **Identification of gene expression profiles that predict the aggressive behavior of breast cancer cells**. *Cancer Res* 2001, **61**(13):5168-5178.
100. Ramaswamy S, Ross KN, Lander ES, Golub TR: **A molecular signature of metastasis in primary solid tumors**. *Nat Genet* 2003, **33**(1):49-54.
101. Martin KJ, Patrick DR, Bissell MJ, Fournier MV: **Prognostic breast cancer signature identified from 3D culture model accurately predicts clinical outcome across independent datasets**. *PLoS One* 2008, **3**(8):e2994.
102. Dai H, van't Veer L, Lamb J, He YD, Mao M, Fine BM, Bernards R, van de Vijver M, Deutsch P, Sachs A *et al*: **A cell proliferation signature is a marker of extremely poor outcome in a subpopulation of breast cancer patients**. *Cancer Res* 2005, **65**(10):4059-4066.
103. Sotiriou C, Pusztai L: **Gene-expression signatures in breast cancer**. *N Engl J Med* 2009, **360**(8):790-800.
104. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT *et al*: **Gene expression profiling predicts clinical outcome of breast cancer**. *Nature* 2002, **415**(6871):530-536.
105. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ *et al*: **A gene-expression signature as a predictor of survival in breast cancer**. *N Engl J Med* 2002, **347**(25):1999-2009.
106. Glas AM, Floore A, Delahaye LJ, Witteveen AT, Pover RC, Bakx N, Lahti-Domenici JS, Bruinsma TJ, Warmoes MO, Bernards R *et al*: **Converting a breast cancer microarray signature into a high-throughput diagnostic test**. *BMC Genomics* 2006, **7**:278.
107. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, d'Assignies MS, Bergh J, Lidereau R, Ellis P *et al*: **Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer**. *J Natl Cancer Inst* 2006, **98**(17):1183-1192.
108. Cardoso F, Van't Veer L, Rutgers E, Loi S, Mook S, Piccart-Gebhart MJ: **Clinical Application of the 70-Gene Profile: The MINDACT Trial**. *J Clin Oncol* 2008, **26**(5):729-735.
109. http://www.eortc.be. In.; 2009.

110. **http://www.oncotypedx.com/** [http://www.oncotypedx.com/]
111. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J *et al*: **Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer**. *J Clin Oncol* 2006, **24**(23):3726-3734.
112. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T *et al*: **A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer**. *N Engl J Med* 2004, **351**(27):2817-2826.
113. Habel LA, Shak S, Jacobs MK, Capra A, Alexander C, Pho M, Baker J, Walker M, Watson D, Hackett J *et al*: **A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients**. *Breast Cancer Res* 2006, **8**(3):R25.
114. Beck AH, Espinosa I, Gilks CB, van de Rijn M, West RB: **The fibromatosis signature defines a robust stromal response in breast carcinoma**. *Lab Invest* 2008, **88**(6):591-601.
115. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, Joshi MB, Harpole D, Lancaster JM, Berchuck A *et al*: **Oncogenic pathway signatures in human cancers as a guide to targeted therapies**. *Nature* 2006, **439**(7074):353-357.
116. Chang HY, Nuyten DS, Sneddon JB, Hastie T, Tibshirani R, Sorlie T, Dai H, He YD, van't Veer LJ, Bartelink H *et al*: **Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival**. *Proc Natl Acad Sci U S A* 2005, **102**(10):3738-3743.
117. Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K, Chi JT, van de Rijn M, Botstein D, Brown PO: **Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds**. *PLoS Biol* 2004, **2**(2):E7.
118. Chen JL, Lucas JE, Schroeder T, Mori S, Wu J, Nevins J, Dewhirst M, West M, Chi JT: **The genomic analysis of lactic acidosis and acidosis response in human cancers**. *PLoS Genet* 2008, **4**(12):e1000293.
119. Chi JT, Wang Z, Nuyten DS, Rodriguez EH, Schaner ME, Salim A, Wang Y, Kristensen GB, Helland A, Borresen-Dale AL *et al*: **Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers**. *PLoS Med* 2006, **3**(3):e47.
120. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, Chen H, Omeroglu G, Meterissian S, Omeroglu A *et al*: **Stromal gene expression predicts clinical outcome in breast cancer**. *Nat Med* 2008, **14**(5):518-527.
121. Massague J: **Sorting out breast-cancer gene signatures**. *N Engl J Med* 2007, **356**(3):294-297.
122. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J *et al*: **Gene-expression**

profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005, **365**(9460):671-679.
123. West RB, Nuyten DS, Subramanian S, Nielsen TO, Corless CL, Rubin BP, Montgomery K, Zhu S, Patel R, Hernandez-Boussard T *et al*: **Determination of stromal signatures in breast carcinoma**. *PLoS Biol* 2005, **3**(6):e187.
124. Ein-Dor L, Kela I, Getz G, Givol D, Domany E: **Outcome signature genes in breast cancer: is there a unique set?** *Bioinformatics* 2005, **21**(2):171-178.
125. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, Nobel AB, van't Veer LJ, Perou CM: **Concordance among gene-expression-based predictors for breast cancer**. *N Engl J Med* 2006, **355**(6):560-569.
126. Ma XJ, Wang Z, Ryan PD, Isakoff SJ, Barmettler A, Fuller A, Muir B, Mohapatra G, Salunga R, Tuggle JT *et al*: **A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen**. *Cancer Cell* 2004, **5**(6):607-616.
127. Gruvberger SK, Ringner M, Eden P, Borg A, Ferno M, Peterson C, Meltzer PS: **Expression profiling to predict outcome in breast cancer: the influence of sample selection**. *Breast Cancer Res* 2003, **5**(1):23-26.
128. Dvorak HF: **Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing**. *N Engl J Med* 1986, **315**(26):1650-1659.
129. Eifel P, Axelson JA, Costa J, Crowley J, Curran WJ, Jr., Deshler A, Fulton S, Hendricks CB, Kemeny M, Kornblith AB *et al*: **National Institutes of Health Consensus Development Conference Statement: adjuvant therapy for breast cancer, November 1-3, 2000**. *J Natl Cancer Inst* 2001, **93**(13):979-989.
130. Goldhirsch A, Glick JH, Gelber RD, Senn HJ: **Meeting highlights: International Consensus Panel on the Treatment of Primary Breast Cancer**. *J Natl Cancer Inst* 1998, **90**(21):1601-1608.
131. Smith MCP, Luker KE, Garbow JR, Prior JL, Jackson E, Piwnica-Worms D, Luker GD: **CXCR4 regulates growth of both primary and metastatic breast cancer**. *Cancer research* 2004, **64**(23):8604.
132. Singer CF, Gschwantler-Kaulich D, Fink-Retter A, Haas C, Hudelist G, Czerwenka K, Kubista E: **Differential gene expression profile in breast cancer-derived stromal fibroblasts**. *Breast Cancer Res Treat* 2008, **110**(2):273-281.
133. Casey T, Bond J, Tighe S, Hunter T, Lintault L, Patel O, Eneman J, Crocker A, White J, Tessitore J *et al*: **Molecular signatures suggest a major role for stromal cells in development of invasive breast cancer**. *Breast Cancer Res Treat* 2009, **114**(1):47-62.
134. Rinn JL, Bondre C, Gladstone HB, Brown PO, Chang HY: **Anatomic demarcation by positional variation in fibroblast gene expression programs**. *PLoS Genet* 2006, **2**(7):e119.

135. Chang HY, Chi JT, Dudoit S, Bondre C, van de Rijn M, Botstein D, Brown PO: **Diversity, topographic differentiation, and positional memory in human fibroblasts**. *Proc Natl Acad Sci U S A* 2002, **99**(20):12877-12882.
136. Sato N, Maehara N, Goggins M: **Gene expression profiling of tumor-stromal interactions between pancreatic cancer cells and stromal fibroblasts**. *Cancer Res* 2004, **64**(19):6950-6956.
137. Roepman P, de Koning E, van Leenen D, de Weger RA, Kummer JA, Slootweg PJ, Holstege FC: **Dissection of a metastatic gene expression signature into distinct components**. *Genome Biol* 2006, **7**(12):R117.
138. Roepman P, Wessels LF, Kettelarij N, Kemmeren P, Miles AJ, Lijnzaad P, Tilanus MG, Koole R, Hordijk GJ, van der Vliet PC *et al*: **An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas**. *Nat Genet* 2005, **37**(2):182-186.
139. Bryant PA, Smyth GK, Robins-Browne R, Curtis N: **Detection of gene expression in an individual cell type within a cell mixture using microarray analysis**. *PLoS One* 2009, **4**(2):e4427.
140. Akiri G, Sabo E, Dafni H, Vadasz Z, Kartvelishvily Y, Gan N, Kessler O, Cohen T, Resnick M, Neeman M *et al*: **Lysyl oxidase-related protein-1 promotes tumor fibrosis and tumor progression in vivo**. *Cancer Res* 2003, **63**(7):1657-1666.
141. Ruiz C, Huang W, Hegi ME, Lange K, Hamou MF, Fluri E, Oakeley EJ, Chiquet-Ehrismann R, Orend G: **Growth promoting signaling by tenascin-C [corrected]**. *Cancer Res* 2004, **64**(20):7377-7385.
142. Sasaki H, Yu CY, Dai M, Tam C, Loda M, Auclair D, Chen LB, Elias A: **Elevated serum periostin levels in patients with bone metastases from breast but not lung cancer**. *Breast Cancer Res Treat* 2003, **77**(3):245-252.
143. Shao R, Bao S, Bai X, Blanchette C, Anderson RM, Dang T, Gishizky ML, Marks JR, Wang XF: **Acquired expression of periostin by human breast cancers promotes tumor angiogenesis through up-regulation of vascular endothelial growth factor receptor 2 expression**. *Mol Cell Biol* 2004, **24**(9):3992-4003.
144. Jones JI, Clemmons DR: **Insulin-like growth factors and their binding proteins: biological actions**. *Endocr Rev* 1995, **16**(1):3-34.
145. Surmacz E: **Function of the IGF-I Receptor in Breast Cancer**. *Journal of Mammary Gland Biology and Neoplasia* 2000, **5**(1):95-105.
146. Hewish M, Chau I, Cunningham D: **Insulin-like growth factor 1 receptor targeted therapeutics: novel compounds and novel treatment strategies for cancer medicine**. *Recent Pat Anticancer Drug Discov* 2009, **4**(1):54-72.
147. Sorrell JM, Baber MA, Caplan AI: **A Self-Assembled Fibroblast-Endothelial Cell Co-Culture System That Supports in vitro Vasculogenesis by both Human Umbilical Vein Endothelial Cells and Human Dermal Microvascular Endothelial Cells**. *Cells Tissues Organs* 2007, **186**(3):157-168.

148. Yauch RL, Gould SE, Scales SJ, Tang T, Tian H, Ahn CP, Marshall D, Fu L, Januario T, Kallop D *et al*: **A paracrine requirement for hedgehog signalling in cancer**. *Nature* 2008, **455**(7211):406-410.
149. Deb S, Amin S, Imir AG, Yilmaz MB, Suzuki T, Sasano H, Bulun SE: **Estrogen Regulates Expression of Tumor Necrosis Factor Receptors in Breast Adipose Fibroblasts**. *J Clin Endocrinol Metab* 2004, **89**(8):4018-4024.
150. Souaze F, Dupouy S, Viardot-Foucault V, Bruyneel E, Attoub S, Gespach C, Gompel A, Forgez P: **Expression of Neurotensin and NT1 Receptor in Human Breast Cancer: A Potential Role in Tumor Progression**. *Cancer Res* 2006, **66**(12):6243-6249.
151. Wang Z, Hao Y, Lowe AW: **The Adenocarcinoma-Associated Antigen, AGR2, Promotes Tumor Growth, Cell Migration, and Cellular Transformation**. *Cancer Res* 2008, **68**(2):492-497.
152. Curran T, Ng JM: **Cancer: Hedgehog's other great trick**. *Nature* 2008, **455**(7211):293-294.
153. Dunn SE, Hardman RA, Kari FW, Barrett JC: **Insulin-like growth factor 1 (IGF-1) alters drug sensitivity of HBL100 human breast cancer cells by inhibition of apoptosis induced by diverse anticancer drugs**. *Cancer Res* 1997, **57**(13):2687-2693.
154. Gooch JL, Van Den Berg CL, Yee D: **Insulin-like growth factor (IGF)-I rescues breast cancer cells from chemotherapy-induced cell death--proliferative and anti-apoptotic effects**. *Breast Cancer Res Treat* 1999, **56**(1):1-10.
155. Patel AV, Cheng I, Canzian F, Le Marchand L, Thun MJ, Berg CD, Buring J, Calle EE, Chanock S, Clavel-Chapelon F *et al*: **IGF-1, IGFBP-1, and IGFBP-3 polymorphisms predict circulating IGF levels but not breast cancer risk: findings from the Breast and Prostate Cancer Cohort Consortium (BPC3)**. *PLoS One* 2008, **3**(7):e2578.
156. Avril-Sassen S, Goldstein L, Stingl J, Blenkiron C, Le Quesne J, Spiteri I, Karagavriilidou K, Watson C, Tavare S, Miska E *et al*: **Characterisation of microRNA expression in post-natal mouse mammary gland development**. *BMC Genomics* 2009, **10**(1):548.
157. Hu M, Yao J, Cai L, Bachman KE, van den Brule F, Velculescu V, Polyak K: **Distinct epigenetic changes in the stromal cells of breast cancers**. *Nat Genet* 2005, **37**(8):899-905.

MICHAL RAJSKI

Address:	In den Klostermatten 32
	4052 Basel
	Switzerland
Telephone number:	(+41) 762 007 299
E- mail address:	michal.rajski@unibas.ch

Personal data:

Date of birth: 21. November.1981

Nationality: Polish

Marital status: Single

Education:

2006 – 2010 PhD Studies Programme at the UNIVERSITY OF BASEL,
Faculty of Science. Basel, Switzerland.
PhD Thesis: "Tumor-stroma derived gene expression patterns as prognosticators in breast cancer".

2000 – 2005 MSc, MSc Studies Programme at the UNIVERSITY OF LODZ,
Faculty of Biology and Environmental Protection. Lodz, Poland.
Master Thesis: "The role of TNF – ∝ in tuberculosis".

1996 – 2000 IX High school of Lodz. Lodz, Poland. Graduated with matriculation.

Professional experience:

Since July 2006
PhD thesis project. Laboratory of Medical Oncology, Department of Biomedicine, University Hospital Basel. Basel, Switzerland.

July 2005 – July 2006
Annual internship in Genetic Toxicology and Safety Pharmacology Department. Novartis Pharma AG. Basel, Switzerland.

September 2004 – June 2005
Master thesis project. Department of Immunology and Infectious Biology. University of Lodz. Lodz, Poland.

September 2003 – June 2004
Annual internship in Clinical Pathology Department. Clinical Microbiology. Clinical Pathology. Laboratory of Clinical Pathology „Centrum Diagnostyki Laboratoryjnej". Lodz, Poland.

September 2002 – June 2003
Students research project: "Biosorption of heavy metals by fungi". Microbiology – Biotechnology Students` Interest Group. University of Lodz. Lodz, Poland.

September 2001 – June 2002

Students research project: "Degradation of chosen xenobiotics by fungi *Mucor sp*." Microbiology – Biotechnology Students` Interest Group. University of Lodz. Lodz, Poland.

Publications:

- **"Tumor-endothelial interaction links the CD44+/CD24- stem cell signature with poor prognosis in early stage breast cancer" -** M. Buess, **M. Rajski,** B. Vogel, R. Herrmann and Ch. Rochlitz – *Neoplasia*, Vol. 11, No. 10, 2009.
- **"IGF–I induced genes in stromal fibroblasts predict the clinical outcome of breast and lung cancer patients" – M. Rajski**, R. Zanetti, B. Vogel, R. Herrmann, Ch. Rochlitz and M. Buess – manuscript accepted in *BMC Medicine*.

Conference calls:

- **EuCC** (European Cancer Center) 15th annual symposium; 2008, **M. Rajski**, B. Vogel, R. Zanetti, R. Herrmann, Ch. Rochlitz and M. Buess; "IGF-I-induced genes in tumor and stroma cells subdivide human breast cancer in four subgroups with significantly different prognosis"
- **ASCO** (American Society of Clinical Oncology), 45th annual meeting; 2009, M. Buess, **M. Rajski**, B. Vogel, R. Herrmann, C. Rochlitz; "Tumor endothelial interaction, CD44+/CD24- stem cell signature, and prognosis in early-stage breast cancer".
- **WCLC** (World Conference on Lung Cancer), 13th annual meeting; 2009, M. Buess, **M. Rajski,** R. Zanetti, B. Vogel, R. Herrmann, Ch. Rochlitz; "IGF–I induced gene signature in stromal fibroblasts is prognostic in adenocarcinomas of the lung"

Poster presentations:

- **BioValley Life Sciences Week 2007;** "Influence of hormones and growth factors on tumor –stroma interaction."
- **BioValley Life Sciences Week 2009;** "IGF-I induced genes in stromal fibroblasts predict the clinical outcome of breast and lung cancer patients."

Scholarships:

- Rector of the University of Lodz Scholarship 2002-2003.
- Rector of the University of Lodz Scholarship 2003-2004.
- Rector of the University of Lodz Scholarship 2004-2005.

Honours and prizes:

- Congratulation Letters from Rector of the University of Lodz for outstanding results in academic years: 2003-2004, 2004-2005.
- Traineeship at Novartis Pharma AG: Award based on the IAESTE contest.

Additional courses:

- Radiological Safety Course for Researchers. (September 2006 - Department of Biomedicine. Basel).
- GLP Basic Course. The Theory and Practice of GLP (Good Laboratory Practice). (November 2005 – Novartis Pharma A.G. Basel).
- Two years teaching course at the University of Lodz with teaching practice at a primary school in Lodz (2003) and at a high school in Lodz (2004).

Theoretical background:

- Oncology.
- Cell Biology.
- Molecular Biology.
- Bioinformatics.
- Genetics.
- Microbial genetics.
- Immunology.
- Cytobiochemistry.
- Virology.
- Microbiology.
- Medical Microbiology.
- Molecular Basis of Human Diseases
- Biotechnology.
- Organic and Inorganic Chemistry.
- Biostatistics.

Practical Skills:

- Microarrays.
- Microarray analysis.
- Flow Cytometry (FACS).
- FACS data analysis (FlowJo).
- Biostatistical programming (R).
- Cell culture techniques.
- Molecular biology techniques (PCR, Western Blot, Gene Cloning).
- Genetic toxicology assays (Micronucleus test, Comet assay).
- ELISA, Immunoelectrophoresis, Immunoenzymatic Techniques.
- Basic cell biology assays.
- Virus cultivation techniques.
- Microorganisms cultivation and identification techniques.
- Light and confocal microscopy.
- Biostatistics analyses (Prism, R).

Foreign language knowledge:

- English: fluent.
- Polish: mother tongue.
- German: good.

Additional information:

- Computer literate (different OSs)
- Statistical analysis software knowledge: (R and Prism).
- Other interests include alpinism, climbing.
- Image processing software knowledge (Adobe Photoshop, Gimp).
- Basic web design (html).